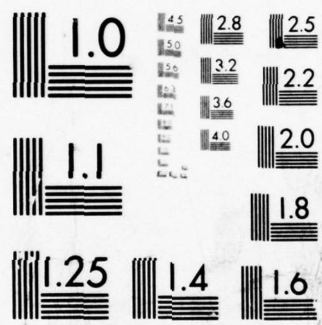
AD-A070 972 KENTUCKY UNIV LEXINGTON DEPT OF MATHEMATICS F/G 12/1
STOCHASTIC REALIZATION AND INVARIANT DIRECTIONS OF THE MATRIX R--ETC(U)
1979 M PAVON AFOSR-78-3519
UNCLASSIFIED AFOSR-TR-79-0807 NL

| OF |

AD
A070972

END
DATE
FILMED
8 --79
DDC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

STOCHASTIC REALIZATION AND INVARIANT DIRECTIONS OF THE MATRIX RICCATI EQUATION

by

Michele Pavon*

Abstract

Invariant directions of the Riccati difference equation of Kalman filtering are shown to occur in a large class of prediction problems and to be related to a certain invariant subspace of the transpose of the feedback matrix. The discrete time stochastic realization problem is studied in its deterministic as well as probabilistic aspects. In particular a new derivation of the classification of the minimal markovian representations of the given process z is presented which is based on a certain backward filter of the innovations. For each markovian representation which can be determined from z the space of invariant directions is decomposed into two subspaces, one on which it is possible to predict the state process without error forward in time and one on which this can be done backward in time.

*Graduate School of Mathematics, University of Kentucky, Lexington, Kentucky, 40506, U.S.A. This work was partially supported by the National Science Foundation under grants MPS75-07028 and MCS77-00241 and partially by the Air Force Office of Scientific Research under Grant AFOSR-78-3519

Introduction

The aim of this paper is to extend the theory of invariant directions of the matrix Riccati equation to a large class of filtering problems, to present some new results on the deterministic and probabilistic aspects of the discrete time stochastic realization problem and to illustrate the particular features introduced in stochastic realization by the presence of invariant directions.

Part 1 of the paper is concerned with characterizing invariant vectors for the usual linear least squares estimation problem in additive white noise. We extend the previous results on the colored noise problem [8, 14, 29] to our more general setting and present some new ones. The main result of this part is Theorem 1.6 which provides different necessary and sufficient conditions for invariance. These conditions are phrased in terms of the convolution of two weighting patterns, of the optimal control of the dual problem, of the best one step predictor and of the feedback matrix $\Gamma(t)$ of the Kalman filter. The latter characterization appears here for the first time. Indeed, the space of all invariant directions is simply the invariant subspace related to the eigenvalue zero of the transpose of $\Gamma(t)$ for t larger than a certain value. This interpretation turns out to be quite useful and enlightening, since $\Gamma(\cdot)$ is the transition matrix of the estimation error and it is essential in classifying markovian representations in the stochastic realization setting (see e.g., Theorem 2.8). Also the fact that invariant vectors are generalized eigenvectors sheds new light on the proof techniques employed in [8, 9, 29]. The paper [9] by Clements and B. D. O. Anderson, which contains results closely related to conditions (ii) and (iii) of

Theorem 1.6, became available to us right after the first version of this paper was submitted. The emphasis in [9], however, is somewhat different from ours in that the authors seek to characterize invariance for a very general form of the linear quadratic regulator problem, whereas our main interest lies in the stochastic implications of this phenomenon.

The second part of the paper deals with discrete time stochastic realization theory. Given a wide sense stationary vector process z with rational spectral density Φ, such that $\Phi(\infty)$ is finite and $\Phi(e^{i\omega})$ is positive definite for all ω, and a Hilbert space H containing the components of $z(t)$ for all t, consider the problem of determining all minimal markovian representations of z (*stochastic realizations*) driven by a white noise with components in H. We solve the problem in the following way. First the second order properties of the stochastic realizations are described. Our results integrate those of B. D. O. Anderson [3-5], Faurre [11, 12] and Ruckebusch [33, 34]. In particular, we show that the correspondence in [33; p.70] between realizations with square transfer function and real symmetric solutions of a certain algebraic matrix equation of the Riccati type holds without any assumption on the feedback matrix. Our analysis on this aspect of the stochastic realization problem parallels in some respects the continuous time work of Lindquist and Picci [19].

Then we turn to the probabilistic side of the problem which has received considerable attention in recent years [1, 2, 18-23, 27, 32-36]. A tool for this study is provided to us by Theorem 2.5, which establishes a correspondence between the deterministic as well as stochastic elements of realizations evolving forward and backward in time. The last two sections of Part 2 are devoted to a new derivation of the classification of the state processes of stochastic realizations due to Ruckebusch [33] in discrete time and Lindquist and Picci in continuous time [19]. Our approach makes essential use of markovian representations of the innovation process with the estimation error as the state. Ruckebusch has used the error process in finite and infinite dimensional stochastic realization to derive a number of results [33-35], but our idea of associating it with a stochastic realization of the innovations appears to be new. Tackling the problem in this way we not only derive the main results in a rather simple manner, but we also gain insight into their meaning. For instance, the important result that realizations which can be constructed from only the process z (*internal*) are in one to one correspondence with the invariant subspaces of the feedback matrix Γ_* (Theorem 2.8) can be given a natural explanation in terms of the backward filter of the innovations (see Remark 2.10). Last, but not least, these stochastic realizations of the innovation process provide a key to understanding the relationship between the invariant subspaces of Γ_* and a certain class of *inner* functions in terms of which it is possible to describe the realizations of z [21, 35, 36]. Our results on this subject, however, will be presented elsewhere.

Part 3 is the natural continuation of Parts 1 and 2 in that it explores how invariant directions affect the family of stochastic realizations. Indeed the space of invariant vectors I is the same for all realizations and is nontrivial if and only if $\Phi(\infty)$ is singular. The characterization of I as the invariant subspace of the transpose of

Γ_* relative to zero is important in establishing the two principal results of Part 3. The first is Theorem 3.8 which says, loosely speaking, that in an invariant direction we can either predict or smooth the state of an internal realization exactly (i.e., without error), showing that I is closely related to the *germ space* of z [23]. The second is Theorem 3.9 which embeds every internal realization in a chain of internal realizations (totally ordered with respect to state covariances) whose minimum element has a full set of *predictable* directions [14] and whose maximum one has a full set of *smoothable* directions (Definition 3.7).

The last section of Part 3 is devoted to comparing two possible approaches to discrete time stochastic realization based on different factorizations of the covariance operator. We show that the factorization leading to markovian representations without noise in the output [1, 11] considerably narrows, compared with the other approach, the solution class of the stochastic realization problem when $\Phi(\infty)$ is singular. This deficiency of the first method makes it advisable to seek markovian representations of the type considered in this paper unless nonsingularity of $\Phi(\infty)$ is guaranteed.

It is worthwhile remarking that the assumptions made on the process z in Parts 2 and 3 are mostly for simplicity. Indeed many of the central results can be established, in a suitably modified form, in the nonstationary case under mild assumptions on z, albeit the derivation becomes more involved. This explains why we refrain from introducing backward realizations and related concepts, like that of smoothable direction, in the setting of Part 1. Our results on this matter will be presented somewhere else.

The scalar case has some interesting features for which we refer the reader to [23].

Part 1

INVARIANT DIRECTIONS OF THE MATRIX RICCATI EQUATION

1.1 Basic notation and formulation of the problem.

We use standard vector-matrix notation, with the following conventions. The unit matrix is denoted by I, the transpose of a matrix by prime. All vectors without prime are column vectors. $\mathcal{N}(R)$ indicates the null space of the matrix R. If R is symmetric, $R > 0$ $(R \geq 0)$ means R positive (nonnegative) definite. If $R \geq 0$, $R^{1/2}$ is the unique nonnegative square root of R. The Moore-Penrose pseudoinverse [26] is denoted by #. The trace operator is indicated by tr. The cone of symmetric, nonnegative definite $n \times n$ matrices is denoted by C_n. Kronecker symbol is δ_{st}. The superscript o identifies "optimal."

Consider the linear stochastic model

$$(1.1) \qquad x(t+1) = Ax(t) + Bw(t)$$

$$(1.2) \qquad y(t) = Cx(t) + Dw(t)$$

with initial condition $x(0) = x_o$, where A, B. C and D are constant matrices of dimensions $n \times n$, $n \times p$, $m \times n$ and $m \times p$, x_o is an n-dimensional zero-mean random vector, the input w is a p-dimensional zero-mean white noise sequence uncorrelated with x_o, $E\{x_o x_o'\} = P_o$ and $E\{w(s)w(t)'\} = I\delta_{st}$.

As is well-known, the best linear least-squares estimate $\hat{x}(t)$ of $x(t)$, given the data $\{y(0), \ldots, y(t-1)\}$, is generated recursively by the Kalman filter

$$\hat{x}(t+1) = A\hat{x}(t) + K(t)[y(t) - C\hat{x}(t)] \qquad \hat{x}(0) = 0 \tag{1.3}$$

where $K(t)$ is given by

$$K(t) = (A\Sigma(t)C' + BD')(C\Sigma(t)C' + DD')^{\#} \tag{1.4}$$

and $\Sigma(t)$ satisfies the Riccati difference equation

$$\begin{cases} \Sigma(t+1) = A\Sigma(t)A' \\ \qquad\qquad - (A\Sigma(t)C' + BD')(C\Sigma(t)C' + DD')^{\#}(C\Sigma(t)A' + DB') \\ \qquad\qquad + BB' \\ \Sigma(0) = P_o \end{cases} \tag{1.5}$$

We shall indicate the solution of (1.5) at time s by $\Sigma(s; P_o)$ when we intend to emphasize the dependence on the initial condition P_o.

Definition 1.1 ([8]). The n-dimensional vector a is called an *s-invariant direction* of equation (1.5) if $a'\Sigma(t; P_o) = a'\Sigma(s; 0)$ for all $t \geq s$ and all $P_o \in C_n$.

We shall study the problem of characterizing all invariant directions of equation (1.5).

1.2 Preliminaries.

In this section we transcribe some well known results of duality between estimation and control into a form best suited to our problem.

We refer the reader to [24] for the variational principles underlying this duality.

Since $\hat{x}(t+1)$ is in the linear span of $y(0), \ldots, y(t)$ there exist matrices $U(s, t)^{o}$ for $s = 0, \ldots, t$ such that $\hat{x}(t+1) = -\sum_{s=0}^{t} (U(s, t)^{o})' y(s)$. Such sequence is optimal for the following *dual problem*: find $U(t) = (U(0, t), \cdots, U(t, t))$ which minimizes

$$\text{(1.6)}\quad \mathrm{tr}\{J[U(t)]\} = \mathrm{tr}\{Q(-1, t)' P_o Q(-1, t) + \sum_{s=0}^{t} Z(s, t)' Z(s, t)\}$$

where

$$\text{(1.7)}\quad Q(s-1, t) = A'Q(s, t) + C'U(s, t) \qquad Q(t, t) = I$$

$$\text{(1.8)}\quad Z(s, t) = B'Q(s, t) + D'U(s, t)$$

A standard argument yields the closed-loop form of the optimal control

$$\text{(1.9)}\qquad U(s, t)^{o} = -K(s)' Q(s, t)^{o} \qquad s = 0, \ldots, t$$

Consider the linear estimator of $x(t+1)$ given by $\gamma(t+1) = -\sum_{s=0}^{t} U(s, t)' y(s)$. Then it is easily seen that

$$\text{(1.10)}\quad x(t+1) - \gamma(t+1) = Q(-1, t)' x_o + \sum_{s=0}^{t} Z(s, t)' w(s)$$

Introducing the quantities $P(s, t) = E\{x(s)[x(t+1) - \gamma(t+1)]'\}$, $R(s, t) = E\{y(s)[x(t+1) - \gamma(t+1)]'\}$ and applying the operator $E\{\cdot[x(t+1) - \gamma(t+1)]'\}$ to both sides of (1.1)-(1.2) we obtain, in view of (1.10), the following *adjoint system*

$$\text{(1.11)}\quad P(s+1, t) = AP(s, t) + BZ(s, t) \qquad P(0, t) = P_o Q(-1, t)$$

(1.12) $$R(s, t) = CP(s, t) + DZ(s, t)$$

The terminology is justified by the fact that, setting up the discrete minimum principle for the dual problem (1.11) are seen to be, with the appropriate normalization, the adjoint equations. Let us note that

(1.13) $$R(s, t) = 0 \qquad s = 0, \ldots, t$$

is a necessary and sufficient condition for optimality of the U(t) sequence. Whenever A is nonsingular we can rewrite (1.7) in the form

(1.14) $$Q(s, t) = (A')^{-1}Q(s - 1, t) - (A')^{-1}CU(s, t) \qquad Q(t, t) = I$$

Hence we have the following input-output relations:

(1.15) $$Z(s, t) = \sum_{i=0}^{s} \tilde{T}(i)\, U(s - i, t) + B'(A')^{-s-1}Q(-1, t)$$

(1.16) $$R(s, t) = \sum_{i=0}^{s} T(i) Z(s - i, t) + CA^{s}P_{o}Q(-1, t)$$

where the weighting patterns $\tilde{T}(\cdot)$ and $T(\cdot)$ are defined by

(1.17) $$\tilde{T}(i) = \begin{cases} D' - B'(A')^{-1}C' & i = 0 \\ -B'(A')^{-i-1}C' & i > 0 \end{cases}$$

(1.18) $$T(i) = \begin{cases} D & i = 0 \\ CA^{i-1}B & i > 0 \end{cases}$$

Combining (1.14) and (1.15) leads us to the *Hamiltonian system*

$$(1.19)\quad \begin{bmatrix} Q(s, t) \\ P(s+1, t) \end{bmatrix} = \begin{bmatrix} (A')^{-1} & 0 \\ BB'(A')^{-1} & A \end{bmatrix} \begin{bmatrix} Q(s-1, t) \\ P(s, t) \end{bmatrix} + \begin{bmatrix} -(A')^{-1}C \\ BD' - BB'(A')^{-1}C \end{bmatrix} U(s, t);$$

$$\begin{bmatrix} Q(-1, t) \\ P(0, t) \end{bmatrix} = \begin{bmatrix} I \\ P_o \end{bmatrix} Q(-1, t)$$

$$(1.20)\quad R(s, t) = [DB'(A')^{-1}\ C] \begin{bmatrix} Q(s-1, t) \\ P(s, t) \end{bmatrix} + [DD' - DB'(A')^{-1}C']U(s, t)$$

where $Q(-1, t) = (A')^t + \sum_{i=0}^{t} (A')^i C' U(i, t)$. It is clear that the weighting pattern $T_H(\cdot)$ of the Hamiltonian system is just the convolution of $T(\cdot)$ and $\tilde{T}(\cdot)$.

$$(1.21)\quad T_H(i) = [T * \tilde{T}](i) = \sum_{j=0}^{i} T(i-j)\tilde{T}(j)$$

The matrices $T_H(0), \ldots, T_H(n-1)$ will play a central role in establishing necessary and sufficient conditions for invariance.

1.3 Characterization of invariant directions.

We study the case where A is nonsingular. This assumption enables us to derive explicit expressions for the invariant vectors. (The case where no restriction is placed on A and on the definitness of the criterion matrices has been recently investigated in [9]). The three following lemmas extend known results to our more general setting.

Lemma 1.2 *The vector* a *is an* s*-invariant direction of* (1.5) *if and only if*

(1.22) $\quad a \in N(Q(t-s, t)^{o})$ *for all* $t \geq s-1$ *and all* $P_o \in C_n$

Proof. Observe that a control $U(t)$ is optimal for the dual problem if and only if it minimizes $a'J[U(t)]a$ for all $a \in R^n$. The result now follows from a straightforward modification of the argument of Theorem 3 in [29]. //

Notice that optimal quantities in the dual problem depend on the terminal weight P_o. To keep notations simple, we shall refrain from explicitly exhibiting this dependence.

Remark 1.3 The proof of the sufficiency part in Lemma 1.2 relies on the fact that, under condition (1.22), $U(t-i, t)^{o}a$ is invariant over $t \geq s$ for $i = 0, \ldots, s-1$. Moreover, when (1.22) holds, it is easily seen using (1.7)-(1.9) that $a \in N(U(i, t)^{o}) \cap N(Z(i, t)^{o})$ for $i = 0, \ldots, t-s$. In particular it follows from (1.10) that $a'\tilde{x}(t+1) = a' \sum_{i=t-s+1}^{t} (Z(i, t)^{0})' w(i)$, where $\tilde{x}(t) = x(t) - \hat{x}(t)$ is the estimation error.

The mathematical framework set up in the previous section will be useful in proving the following result.

Lemma 1.4 *The vector* a *satisfies* (1.22) *if and only if*

$$a = -\sum_{i=1}^{s} (A')^{-i} C' \lambda_i \tag{1.23}$$

where the m-dimensional vectors $\lambda_1, \lambda_2, \ldots, \lambda_s$ *are such that*

$$\sum_{i=0}^{s-j} T_H(i)\, \lambda_{j+i} = 0 \qquad j = 1, \ldots, s \tag{1.24}$$

In this case the optimal control satisfies

(1.25) $\qquad U(t)^{o}a = (0, \ldots, 0, \lambda_s, \ldots, \lambda_1)$

Proof. Assume that (1.22) holds. In view of the time invariance discussed in Remark 1.3, we can set $\lambda_i = U(t - i + 1, t)^{o}a$ for $i = 1, \ldots, s$. Expression (1.23) can now be derived using (1.7) recursively. Let us consider the input-output relation of the Hamiltonian system

$$R(s, t) = [DB'(A')^{-1} \; C]A_H^s\begin{bmatrix} I \\ P_o \end{bmatrix}Q(-1, t) + \sum_{i=0}^{s} T_H(i)U(s - i, t)$$

where

$$A_H = \begin{bmatrix} (A')^{-1} & 0 \\ BB'(A')^{-1} & A \end{bmatrix}$$

As observed in Remark 1.3, $a \in N(Q(-1, t)^{o})$. Then (1.24) follows from the optimality conditions (1.13). Conversely suppose a is as in (1.23) with the λ_j:s satisfying (1.24). Using (1.9) and, recursively, (1.7), we obtain

$$U(k, t)^{o}a = -K(k)'[(A')^{t-k} + \sum_{i=1}^{t-k} (A')^{i-1}C'U(k + i, t)^{o}]a$$

which, together with (1.23) yields

$$U(k, t)^{o}a = -K(k)'\{-\sum_{i=1}^{s-t+k} (A')^{-i}C'\lambda_{t-k+1}$$

$$+ \sum_{i=1}^{t-k} (A')^{i-1}C'[U(k + i, t)^{o}a - \lambda_{t-k-i+1}]\}$$

A calculation similar to that found in the proof of Theorem 8 in [29], i.e., using (1.4), (1.5) repeatedly and condition (1.24), shows that

$$(1.26) \qquad K(k)' \sum_{i=1}^{s-t+k} (A')^{-i} C' \lambda_{t-k+i} = \lambda_{t-k+1}$$

which, inserted into the previous expression for $U(k, t)^o a$, enables us to derive $U(k, t)^o a = \lambda_{t-k+1}$ for $k = t - s + 1, \ldots, t$ recursively. This and (1.7) yield $Q(t - s, t)^o a = 0$, i.e., condition (1.22). Also (1.25) now follows in view of Remark 1.3. This completes the proof. //

A straightforward extension of the proof of Theorem 8 in [29] establishes the following lemma.

Lemma 1.5 *A vector* a *is* s-*invariant for* (1.5) *if and only if* a *is as in* (1.23) *and*

$$(1.27) \qquad a'\hat{x}(t + 1) = - \sum_{i=1}^{s} \lambda_i' y(t + 1 - i) \quad \text{for all} \quad t \geq s - 1$$

Let $\Gamma(t)$ denote the *feedback matrix* $A - K(t)C$.

Theorem 1.6 *The following statements are equivalent:*

(i) a *is an* s-*invariant direction of* (1.5).

(ii) a *satisfies* (1.22).

(iii) a *is as in* (1.23) *and* (1.24) *holds.*

(iv) a *is as in* (1.23) *and* (1.27) *holds.*

(v) a *generates the same* s-*dimensional cyclic subspace of* $\Gamma(t)'$ *for all* $t \geq s - 1$ *and all* $P_o \in C_n$*; this invariant subspace of* $\Gamma(t)'$ *is associated with the eigenvalue zero, i.e.,* $(\Gamma(t)')^s a = 0$. *Moreover* $\Gamma(t - s + 1)' \cdots \Gamma(t)'a = 0$ *for all* $t \geq s - 1$.

Proof. The equivalence of (i), (ii), (iii) and (iv) follows directly from Lemmas 1.2, 1.4 and 1.5. Suppose a satisfies (v) and observe that relations (1.7) and (1.9) yield the expression $Q(t-s,t)^{o} = \Gamma(t-s+1)' \cdots \Gamma(t)'$. By assumption $\Gamma(t-s+1)' \cdots \Gamma(t)'a = 0$ and (1.22) follows. Conversely, if we assume (iii), we derive from (1.26) and the last part of the proof of Lemma 1.4 the relation

$$\Gamma(t)' \sum_{i=1}^{s-j} (A')^{-i} C' \lambda_{i+j} = \sum_{i=1}^{s-j-1} (A')^{-i} C' \lambda_{i+j+1}$$

for all $t \geq s-j-1$ and all $P_o \in C_n$, where $j = 1, \ldots, s-1$ and, for $j = s-1$, the right hand side is defined to be zero. This establishes (v). //

Condition (v) of this theorem is new. Its importance will completely surface in the stochastic realization setting.

Remark 1.7 ([8]). The sets I_s of s-invariant directions and $I = \cup_{s=1}^{\infty} I_s$ of invariant directions are vector spaces. It follows from the previous theorem that $I = \cup_{s=1}^{n} I_s$.

Remark 1.8. The dimension of the *invariant subspace* I can be easily determined in the single-output case $y(t) = c'x(t) + d'w(t)$. It is equal to the minimum between the rank of the observability matrix $[c \;\; A'c \cdots (A')^{n-1}c]'$ and the first index j such that $T_H(j-1) = \cdots = T_H(0) = 0$ and $T_H(j) \neq 0$. The general case is rather involved. We shall not pursue here the extension of the results of [29] on this matter.

Let

(1.28) $$W(z) = \sum_{i=0}^{\infty} T(i)z^{-i} = C(zI - A)^{-1}B + D$$

be the *transfer function* of (1.1)-(1.2) and

(1.29) $$W_H(z) = \sum_{i=0}^{\infty} T_H(i)z^{-i}$$

the transfer function of the Hamiltonian system. The following characterization of $T_H(\cdot)$ will be helpful in the third part of the paper.

Theorem 1.9 *Assume* A *nonsingular. Then*

(1.30) $$W_H(z) = W(z)W(z^{-1})'$$

If y *in* (1.2) *is stationary with spectral density* $\Phi(z)$, *we also have*

(1.31) $$W_H(z) = \Phi(z)$$

Proof. Consider $W(z^{-1})' = B'(z^{-1}I - A')^{-1}C' + D' =$ $-B'(A')^{-1}(I - z^{-1}(A')^{-1})^{-1}C' + D'$. Expand the last term in a neighborhood of infinity as follows:

(1.32) $$\begin{aligned} &-B'(A')^{-1}(I - z^{-1}(A')^{-1})^{-1}C' + D' \\ &= D' - B'(A')^{-1}C' - B'(A')^{-2}C'z^{-1} - B'(A')^{-3}C'z^{-2} \dots \\ &= \sum_{i=0}^{\infty} \tilde{T}(i)z^{-i} \end{aligned}$$

Take the Cauchy product of the two series in (1.28) and (1.32) to get (1.30). In the case of a stationary y the well-known spectral factorization formula

$$\Phi(z) = W(z)W(z^{-1})' \tag{1.33}$$

yields (1.31). //

Notice that the calculations in the previous theorem make sense because the series in (1.28) and (1.29) converge respectively to $W(z)$ and to $W_H(z)$ in an appropriate neighborhood of infinity.

Let $\Delta(t, s) = E\{y(t)y(s)'\}$ be the covariance operator of the observations. It is a simple matter, using the expression $y(s) = CA^{-n} x(s + n) + \sum_{i=0}^{n-1} \tilde{T}(i)'w(s + i)$ which can be derived from (1.1)-(1.2), to see that the parameters $T_H(0), \cdots, T_H(n - 1)$ determine the degree of "smoothness" of $\Delta(\cdot,\cdot)$, i.e., the number of differencing operations on $\Delta(\cdot,\cdot)$ necessary in each direction to produce a Kronecker delta. This number has been named in the scalar case *relative order of the covariance*, see [14] for example. This fact has its counterpart in the spectral domain in Theorem 1.9.

1.4 Predictable directions.

The invariance properties of invariant directions have been pointed out by several authors [8, 14]. Indeed, as it is apparent from Theorem 1.6, the space I is invariant over models (1.1)-(1.2) having the same covariance of the output and the same (up to a change of basis in the state space) pair (A, C). However, if a is an s-invariant vector for (1.5) the value $a' \Sigma(s; P_o)$ does depend on the model. A special case of particular interest is when $a \in N(\Sigma(s; P_o))$.

Definition 1.10 ([14]). The n-dimensional vector a is called an *s-predictable direction* of equation (1.5) if $a'\Sigma(t; P_o) = a'\Sigma(s; P_o) = 0$ for all $t \geq s$. The two following theorems extend some results of Gevers [14].

Theorem 1.11 *The vector* a *is an* s-*predictable direction of* (1.5) *if and only if* a *is as in* (1.23) *with the* λ_i *satisfying*

$$\text{(1.34)} \qquad \sum_{i=0}^{s-j} \tilde{T}(i)\lambda_{j+i} = 0 \qquad j = 1, \ldots, s$$

Proof. If a is s-predictable $a'\tilde{x}(t + 1) = 0$ for all $t \geq s - 1$. Using (1.10) with optimal quantities we see that $a \in N(Q(-1, t)^o)$ and $a \in \cap_{i=0,\ldots,t} N(Z(i, t)^o)$ for all $t \geq s - 1$. Again time invariance of the optimal control can be shown to hold and, identifying quantities as in (1.25), we get (1.23) from $Q(-1, s - 1)^o a = 0$. Also (1.34) follows from (1.15). To prove the converse first observe that (1.34) implies (1.24). By Lemma 1.4 $a \in N(Q(-1, t)^o)$ and (1.25) holds. From (1.15) and (1.10) we conclude that $a'\tilde{x}(t + 1) = 0$ for all $t \geq s - 1$, i.e., a is s-predictable. //

Theorem 1.12 *Let* $\Sigma(s; P_o) > 0$. *Then* $\Sigma(t; P_o) > 0$ *for all* $t \geq s$ *if and only if* $\tilde{T}(0)$ *has rank* m.

Proof. Let λ be such that $\tilde{T}(0)\lambda = 0$. Then $(A')^{-1}C'\lambda \in N(\Sigma(t; P_o))$ for all $t \geq 1$. To prove the other half we use induction. Suppose $\Sigma(t - 1; P_o) > 0$ and $a \in N(\Sigma(t; P_o))$. It follows from the principle of optimality that

$$(1.35)\quad 0 = a'\,\Sigma(t;\,P_o)a = \min_{\lambda\in R^m}\{(a'A + \lambda'C)\Sigma(t-1;\,P_o)(A'a + C'\lambda) + (a'B + \lambda'D)(B'a + D'\lambda)$$

Let λ^o be the optimal value in (1.35). Since $\Sigma(t-1;\,P_o) > 0$ we get $a = -(A')^{-1}C'\lambda^o$, $B'a + D'\lambda^o = 0$ *and finally* $(D' - B'(A')^{-1}C')\lambda^o = 0$. If $\tilde{T}(0)$ has rank m this implies that $a = 0$. //

Remark 1.13. Theorem 1.12 agrees with the results obtained by Silverman *et al.* [25, 30, 38]. In fact, the presence of nontrivial predictable directions of (1.5) implies that the system (1.1)-(1.2) is not *strongly observable* [38]. However, it can well happen that it is completely observable (and controllable). In the third part of the paper we shall study a set of minimal realizations with a nontrivial invariant and, for some of them, predictable subspace.

1.5 Discussion.

Our study has shown that invariant directions can occur in a more general situation than just the noise-free measurements case treated in [8, 14, 29]. Conditions (iv) and (v) of Theorem 1.6 provide us with a probabilistic interpretation of this phenomenon. In an invariant direction the optimal filter depends only on some of the last observation instead of the whole information available. This fact is strictly related to the invariant subspace of $\Gamma(t)'$ corresponding to zero. Moreover, in the case when y is stationary with rational spectral density, condition (iii) of Theorem 1.6 with Theorem 1.9 shows a precise connection between invariant vectors and the spectrum of y. All of this motivates the stochastic realization approach to the problem taken in Part 3.

Finally we remark that this theory can be extended in a straightforward manner to the case when the system matrices are time-varying replacing the concept of invariant direction by that of *degenerate* direction [14]. A reduction of the order of the Riccati equation which has to be solved can be achieved along the lines of [8] whenever invariant (or degenerate) directions exist.

Part 2

DISCRETE TIME STOCHASTIC REALIZATION: GENERAL THEORY

2.1 Notation and problem formulation.

Almost sure equality between random vectors is simply indicated as equality. If $\{\xi(t);\ t \in Z\}$ is a second order vector process defined on the probability space (Ω, F, P) and S a subset of the integers Z, we denote by $H_s(\xi)$ the closed linear hull in $L_2(\Omega, F, P)$ of the components of $\xi(t)$, $t \in S$. We shall write $H(\xi)$, $H_t^-(\xi)$, $H_t^+(\xi)$ and $H(\xi(t))$ instead of $H_Z(\xi)$, $H_{\{z\in Z|z\le t\}}(\xi)$, $H_{\{z\in Z|z\ge t\}}(\xi)$ and $H_{\{t\}}(\xi)$ respectively. Let $\hat{E}\{\cdot|H_s(\xi)\}$ denote the orthogonal projection operator onto $H_s(\xi)$. We abbreviate $\hat{E}\{\cdot|H(\xi(t))\}$ as $\hat{E}\{\cdot|\xi(t)\}$. The process ξ is called a *wide sense vector Markov process* if

$$\hat{E}\{\xi(s)|H_t^-(\xi)\} = \hat{E}\{\xi(s)|\xi(t)\} \qquad \text{for } s \ge t$$

or equivalently

$$\hat{E}\{\xi(s)|H_t^+(\xi)\} = \hat{E}\{\xi(s)|\xi(t)\} \qquad \text{for } s \le t$$

For the sake of brevity we shall use the word "Markov" instead of the expression "wide sense vector Markov."

We shall be concerned with a wide sense stationary, purely nondeterministic, m-dimensional stochastic process $\{z(t);\ t \in Z\}$. The process z, defined on the probability space (Ω, F, P), is assumed to be centered and to have a rational spectral density Φ such that $\Phi(\infty) < \infty$. The finiteness of $\Phi(\infty)$ is essential only in Part 3 and is assumed here for simplicity. The matrix function $\Phi(\cdot)$ enjoys the following properties: each element of Φ is analytic on the unit circle, Φ is discrete parahermitian, i.e., $\Phi(z)' = \Phi(z^{-1})$ and $\Phi(e^{i\omega}) \geq 0$ hermitian for all real ω. In addition we suppose that z is a *minimal process* [31] which, in view of the rationality of its spectral density, is equivalent to $\Phi(e^{i\omega}) > 0$ for all ω. This assumption too is made for convenience and can be removed without impairing the main results of Parts 2 and 3.

In many problems of estimation and optimal control, when given a non Markov process z which models the information flow, it is necessary to resort to an auxiliary Markov process x which makes $\xi(t) = \begin{pmatrix} x(t) \\ z(t-1) \end{pmatrix}$ a Markov process. More precisely we are interested in the following two problems.

I. *Wide sense stochastic realization problem:* Determine, from the knowledge of Φ, all quadruplets $[A, B, C, D]$, with dimension of A minimal, such that the process y, generated by the dynamical system (1.1) - (1.2) driven by an arbitrary normalized white noise w, has the same spectral density Φ as z.

II. *Proper stochastic realization problem:* Let H be a Hilbert space such that $H(z) \subset H \subset L_2(\Omega, F, P)$. Given H and the process z find all quintuplets $[A, B, C, D; w]$, with dimension of A minimal and w a normalized white noise satisfying $H(w) \subset H$, such that $y(t)$, generated by

(1.1) - (1.2) and $z(t)$ are equivalent random vectors for all t.

We shall call a solution to Problem I a *wide sense minimal stochastic realization* and a solution to Problem II a *proper minimal*[1] *stochastic realization*. It is immediate that to each proper stochastic realization there corresponds a (unique) wide sense realization. The converse is false. To attack Problem II we shall choose a route passing through the solution of Problem I, with the intent of deriving some new results along the way. It is good to bear in mind, however, that a direct probabilistic approach to proper stochastic realization is possible and in a sense more natural [18, 20-22, 27, 35, 36].

2.2 Wide sense stochastic realizations.

Our preliminaries on Problem I are based on the important work of B. D. O. Anderson [3-5] and Faurre [11, 12]. Problem I is equivalent to the classical *spectral factorization problem*. Find all *minimal stable spectral factors* of Φ, i.e., all matrices W of real rational functions of minimal McMillan degree [6] and with all their poles inside the unit circle which satisfy (1.33). Indeed, if $[A, B, C, D]$ solves Problem I, then $W(z) = C(zI - A)^{-1}B + D$ is a stable minimal spectral factor of Φ. Conversely, any such W yields a whole class of wide sense stochastic realizations. In fact, using one of the algorithms [16, 39, 41] available in the literature we can compute a minimal [6] realization $[A, B, C, D]$ of W. Then all minimal realizations of W given by

$$(2.1) \qquad [T^{-1}AT, T^{-1}B, CT, D] \qquad T \in GL_{\dim A}(R)$$

[1] From now on we shall leave the word minimal out. All realizations are to be intended to be minimal unless the opposite is explicitly stated.

solve Problem I. In view of this equivalence Problem I can be solved as follows. Express Φ, by means of partial fractions, as

$$\Phi(z) = S(z) + S(z^{-1})' \tag{2.2}$$

where S is a *positive real*[2] and rational function. Let $[F, G, H, J]$ be a minimal realization of S. As observed before, several procedures are known to determine $[F, G, H, J]$ which is unique up to an equivalence such as in (2.1). The following simple lemma allows us to eliminate J in the sequel.

Lemma 2.1 *Let* S *be the positive real function satisfying* (2.2) *and* $[F, G, H, J]$ *a minimal realization of* S. *If* $\dim F = n \geq 1$, *then* F *is nonsingular and* $J + J' = G'(F')^{-1}H' + \Phi(\infty)$.

Proof. Taking limits in (2.2) we see that $\Phi(\infty) = J + J' + \lim_{z\to\infty} G'(z^{-1}I - F')^{-1}H'$, since $S(z) = H(zI - F)^{-1}G + J$. The conclusion now follows from the finiteness of $\Phi(\infty)$ and the minimality of $[F, G, H, J]$.//

To avoid trivialities, we shall assume from now on that z is not a white noise, i.e., $\dim F = n \geq 1$. It follows from Lemma 2.1 and the celebrated Positive Real Lemma (see e.g., [28]) that the set of all wide sense stochastic realizations is nonempty and given by

$$[A, B, C, D] = [T^{-1}FT,\ T^{-1}(B_1, B_2)V,\ HT,\ (R(P)^{1/2}, 0)V] \tag{2.3}$$

where $T \in GL_n(R)$, V is any $p \times p$ constant orthogonal matrix, B_1 is $n \times m$, B_2 is $n \times (p - m)$ (here $p \geq m$ is arbitrary), P is $n \times n$,

---

[1] A real rational function with no pole on the unit circle is said to be (discrete) positive real if it has no poles outside the unit circle and $S(e^{i\omega}) + S(e^{-i\omega})' \geq 0$ hermitian for all real ω.

symmetric and positive definite, $R(P)$ is the nonnegative definite quantity $G'(F')^{-1}H' + \Phi(\infty) - HPH'$ and (P, B_1, B_2) solve the system

$$(2.4) \qquad P = FPF' + B_1B_1' + B_2B_2'$$

$$(2.5) \qquad G = FPH' + B_1R(P)^{1/2}$$

It is no restriction to choose $T = I$ and $V = I$ in (2.3). In fact all other realizations can be obtained from realizations of the form

$$(2.6) \qquad x(t+1) = Fx(t) + B_1u(t) + B_2v(t) \qquad w = \begin{pmatrix} u \\ v \end{pmatrix}$$

$$(2.7) \qquad z(t) = Hx(t) + R(P)^{1/2}u(t)$$

by means of a change of basis and an orthogonal transformation of w. Hence, whenever convenient, we shall narrow our attention to realizations of the type (2.6) - (2.7). We shall write $\mathcal{P}$ for the set of all symmetric, positive definite P which solve (2.4) - (2.5) and $\mathcal{Q}$ for the subset of $\mathcal{P}$ consisting of those P such that $R(P)$ is singular. Notice that the realizations corresponding to elements of $\mathcal{Q}$ are precisely those which have singular intensity of the noise in the output equation. It can be shown [12] that $\mathcal{P}$ is compact, convex and forms a complete lattice when endowed with the natural partial order $P_1 \geq P_2$ if and only if $P_1 - P_2 \geq 0$. There exist a maximal and a minimal element P^* and P_* so that $P_* \leq P \leq P^*$ for all $P \in \mathcal{P}$. Moreover the minimality of the process z implies [13] that $P^* - P_*$ and $R(P_*)$ are positive definite. Hence $\mathcal{P} \backslash \mathcal{Q}$ is nonempty. The following result provides us with some information about the set $\mathcal{Q}$.

Proposition 2.2 *The set* $\mathcal{P}\backslash\mathcal{Q}$ *is convex. For all* $P \in \mathcal{P}\backslash\mathcal{Q}$, $Q \in \mathcal{Q}$ *and* $\lambda \in (0, 1]$ *we have that* $[\lambda P + (1 - \lambda)Q] \in \mathcal{P}\backslash\mathcal{Q}$. *The set* $\mathcal{Q}$ *is contained in the relative boundary of* $\mathcal{P}$.

Proof. The first two results follow at once from the fact that for P_1, $P_2 \in \mathcal{P}$, $\lambda \in [0, 1]$ we have $R(\lambda P_1 + (1 - \lambda)P_2) = \lambda R(P_1) + (1 - \lambda)R(P_2)$. They in turn imply that, if $P \in \mathcal{P}\backslash\mathcal{Q}$ and $Q \in \mathcal{Q}$, the segment $[P, Q]$ cannot be extended beyond Q without leaving $\mathcal{P}$. We conclude that Q belongs to the relative boundary of $\mathcal{P}$. //

Let us introduce the mapping $\Lambda: R^{n\times n} \longrightarrow R^{n\times n}$ defined by

$$\Lambda(P) = -P + FPF' + (G - FPH')R(P)^{-1}(G' - HPF') \tag{2.8}$$

The set $\mathcal{P}\backslash\mathcal{Q}$ is contained in the domain of $\Lambda(\cdot)$. It is possible to extend $\Lambda(\cdot)$ to all of $\mathcal{P}$ since the points in $\mathcal{Q}$ constitute removable discontinuities. We can now derive an important alternative characterization of the set $\mathcal{P}$.

Theorem 2.3 *Let* $\Lambda(\cdot)$ *be given by* (2.8). *Then* $\mathcal{P} = \{P | P = P', \Lambda(P) \leq 0\}$.

Proof. Let (P, B_1, B_2) solve (2.5)-(2.6) with $P = P'$ and $P > 0$. Then if $P \in \mathcal{P}\backslash\mathcal{Q}$, we get immediately $\Lambda(P) = -B_2B_2'$. If $P \in \mathcal{Q}$, let $\{P_i\}_{i=1}^{\infty}$ be a sequence in $\mathcal{P}\backslash\mathcal{Q}$ converging to P. Then $\Lambda(P_i) \leq 0$ and it follows that $\Lambda(P) = \lim_i \Lambda(P_i) \leq 0$. This shows that $\mathcal{P} \subseteq \{P | P = P', \Lambda(P) \leq 0\}$. The other inclusion can be proven by an argument akin to that used by B. D. O. Anderson [4; p.140]. //

This result provides a bridge between the theory of positive real functions and the study of quadratic matrix inequalities and algebraic Riccati equations.

Let us introduce the set $\mathcal{P}_o = \{P \in \mathcal{P} | \Lambda(P) = 0\}$. Clearly $\mathcal{P}_o$ consists of all $P \in \mathcal{P}$ for which $B_2 = 0$.

Remark 2.4 Since the eigenvalues of F lie in the open unit disc, elementary Liapunov theory ensures that to each (B_1, B_2) there corresponds a unique P. The converse does not hold in general. However, for realizations of the form (2.6)-(2.7), to each P there corresponds a unique B_1. This is immediate from (2.5) for $P \in \mathcal{P} \backslash \mathcal{Q}$ and holds for all $P \in \mathcal{P}$ since points in $\mathcal{Q}$ appear as removable discontinuities of the map $P \to (G - FPH')R(P)^{-1/2}$. Hence there is a unique wide sense realization of the type (2.6)-(2.7) corresponding to each P in $\mathcal{P}_o$.

Both Problem I and II seek to find dynamical systems evolving forward in time like (1.1)-(1.2) which is natural to call *forward representations* of the process z. Yet, there are other representations of interest. There exist situations, for example, in which it is more useful to consider a *backward representation* of the form

$$\bar{x}(t - 1) = \bar{A}\bar{x}(t) + \bar{B}\bar{w}(t) \tag{2.9}$$

$$y(t) = \bar{C}\bar{x}(t) + \bar{D}\bar{w}(t) \tag{2.10}$$

where $\bar{w}$ is a normalized white noise such that $\bar{w}(t)$ is orthogonal to $H_t^+(\bar{x})$ for all t. This leads us to formulate the backward counterpart of Problems I and II.

$\bar{\text{I}}$. *Wide sense backward stochastic realization problem:* Determine, from the knowledge of Φ, all quadruplets $[\bar{A}, \bar{B}, \bar{C}, \bar{D}]$, with dimension of $\bar{A}$ minimal, such that the process y, generated by the dynamical system

(2.9)-(2.10) driven by an arbitrary normalized white noise $\bar{w}$, has the same spectral density Φ as z.

$\overline{\text{II}}$. *Proper backward stochastic realization problem:* Given H and z find all quintuplets $[\bar{A}, \bar{B}, \bar{C}, \bar{D}; \bar{w}]$, with dimension of $\bar{A}$ minimal and $\bar{w}$ a normalized white noise satisfying $H(\bar{w}) \subset H$, such that $y(t)$ given by (2.9)-(2.10) and $z(t)$ are equivalent random vectors for all t.

Solutions to Problems $\overline{\text{I}}$ and $\overline{\text{II}}$ are called *wide sense* and *proper backward stochastic realizations* respectively. We shall now briefly discuss Problem $\overline{\text{I}}$, while Problem $\overline{\text{II}}$ will be implicitly solved in the next three sections in view of Theorem 2.5 below.

Problem $\overline{\text{I}}$ is equivalent to the *dual spectral factorization problem* considered by Anderson [3] and Faurre [12] which consists in finding all minimal unstable (i.e., with all the poles outside the unit circle) spectral factors $\bar{W}(z)$ of $\Phi(z)$. It follows from the parahermitian property of Φ that this problem is equivalent to the spectral factorization problem for $\Phi(\cdot)'$. Hence all the results on Problem $\overline{\text{I}}$ have a natural counterpart in the backward setting via the duality relation $(F, G, H, \Phi(\infty)) \rightarrow (F', H', G', \Phi(\infty)')$. In particular all solutions to Problem $\overline{\text{I}}$ are characterized by

$$(2.11) \qquad [\bar{A}, \bar{B}, \bar{C}, \bar{D}] = [T^{-1}F'T,\ T^{-1}(\bar{B}_1, \bar{B}_2)V,\ G'T,\ (\bar{R}(\bar{P}))^{1/2}, 0)V]$$

where T and V are as in (2.3), $\bar{B}_1$ is $n \times m$, $\bar{B}_2$ is $n \times (p - m)$, $\bar{P}$ is $n \times n$, symmetric and positive definite, $\bar{R}(\bar{P}) = HF^{-1}G + \Phi(\infty)' - G'\bar{P}G$ and $(\bar{P}, \bar{B}_1, \bar{B}_2)$ solve the system

$$\overline{P} = F'\overline{P}F + \overline{B}_1\overline{B}_1' + \overline{B}_2\overline{B}_2' \tag{2.12}$$

$$H' = F'\overline{P}G + \overline{B}_1\overline{R}(\overline{P})^{1/2} \tag{2.13}$$

Whenever it is appropriate, we shall restrict ourselves to realizations of the type

$$\overline{x}(t-1) = F'\overline{x}(t) + \overline{B}_1\overline{u}(t) + \overline{B}_2\overline{v}(t) \qquad \overline{w} = \begin{bmatrix} \overline{u} \\ \overline{v} \end{bmatrix} \tag{2.14}$$

$$z(t) = G'\overline{x}(t) + \overline{R}(\overline{P})^{1/2}\overline{u}(t) \tag{2.15}$$

where $\overline{P}$ is the state covariance. The set $\overline{\mathcal{P}}$ of all symmetric, positive definite solutions to (2.12)-(2.13) and $\overline{\mathcal{Q}}$ of all $\overline{P} \in \overline{\mathcal{P}}$ such that $\overline{R}(\overline{P})$ is singular enjoy the same kind of properties as $\mathcal{P}$ and $\mathcal{Q}$ respectively. In particular there exist $\overline{P}_*$ and $\overline{P}^*$ such that $\overline{P}_* \le \overline{P} \le \overline{P}^*$ for all $\overline{P} \in \overline{\mathcal{P}}$. It is well known [12, 37] that $\overline{\mathcal{P}} = \{P^{-1} | P \in \mathcal{P}\}$, so that $\overline{P}_* = (P^*)^{-1}$ and $\overline{P}^* = (P_*)^{-1}$. Indeed, the following result holds.

Proposition 2.5 *The quadruplet* $[F, B, H, D]$ *with* $B = (B_1, B_2)$ *and* $D = (R(P)^{1/2}, 0)$ *solves Problem I if and only if* $[F', \overline{B}, G', \overline{D}]$ *solves Problem* $\overline{I}$ *where*

$$\overline{B} = (\overline{B}_1, \overline{B}_2) = -P^{-1}F^{-1}B(I - B'P^{-1}B)^{1/2} \tag{2.16}$$

$$\overline{D} = (D - HF^{-1}B)(I - B'P^{-1}B)^{1/2} = \tag{2.17}$$

$$= (R(P)^{1/2} - HF^{-1}B_1, - HF^{-1}B_2)(I - B'P^{-1}B)^{1/2}$$

Proof. The result follows from long but simple calculations using (2.4)-(2.5) and (2.12)-(2.13). //

This proposition exhibits a correspondence between forward and backward wide sense realizations and raises the question whether a result of the same type can be established for proper realizations. We turn to this problem in the beginning of the next section.

2.3 Proper stochastic realizations.

Let us consider a proper stochastic realization of z [F, B, H, D; w], with state process x and state covariance P. As is well known, the orthogonal decomposition

$$x(t+1) = \hat{E}\{x(t+1)|H_t^-(x)\} + [x(t+1) - \hat{E}\{x(t+1)\ H_t^-(x)\}]$$

yields (2.6). Similarly the expression

$$x(t) = \hat{E}\{x(t)|H_{t+1}^+(x)\} + [x(t) - \hat{E}\{x(t)|H_{t+1}^+(x)\}] \tag{2.18}$$

leads to a backward model. In fact, the process x is Markov in both directions and

$$\begin{aligned}\hat{E}\{x(t)|x(t+1)\} &= E\{x(t)x(t)'F'\}E\{x(t+1)x(t+1)'\}^{-1}x(t+1)\\ &= PF'P^{-1}x(t+1)\end{aligned}$$

which gives

$$\begin{aligned}P^{-1}x(t) &= F'P^{-1}x(t+1) - F'P^{-1}B[w(t) - B'(F')^{-1}P^{-1}x(t)]\\ &= F'P^{-1}x(t+1) - P^{-1}F^{-1}B[I - B'P^{-1}B][w(t) -\\ &\qquad - B'(F')^{-1}P^{-1}x(t)]\end{aligned}$$

Defining

$$\bar{x}(t) = P^{-1}x(t+1) \tag{2.19}$$

and

$$(2.20) \qquad \bar{w}(t) = (I - B'P^{-1}B)^{1/2}(w(t) - B'(F')^{-1}P^{-1}x(t))$$

We finally obtain

$$(2.21) \qquad \bar{x}(t-1) = F'\bar{x}(t) - P^{-1}F^{-1}B(I - B'P^{-1}B)^{1/2}\bar{w}(t)$$

It is not difficult to check that $\bar{w}$ is a normalized white noise such that $\bar{w}(t)$ is orthogonal to $H_t^+(\bar{x})$ for all t. The forward and backward noises are related as follows

$$(2.22) \qquad (I - B'P^{-1}B)^{1/2}\bar{w}(t) = w(t) - \hat{E}\{w(t)|H_{t+1}^+(x)\}$$

We also have

$$\begin{aligned} z(t) &= Hx(t) + Dw(t) = [G'(F')^{-1}P^{-1} - DB'(F')^{-1}P^{-1}]x(t) + Dw(t) \\ &= G'P^{-1}x(t+1) + [D - G'P^{-1}B][w(t) - B'(F')^{-1}P^{-1}x(t)] \\ &= G'\bar{x}(t) + [D - HF^{-1}B][I - B'P^{-1}B]^{1/2}\bar{w}(t) \end{aligned}$$

Summing up we obtain a strict sense version of Proposition 2.5, analogous to the continuous time result of Lindquist and Picci [19].

Theorem 2.5 *The quintuplet* $[F, B, H, D; w]$ *is a proper (forward) stochastic realization of* z *with state process* x *and state covariance* P *if and only if the quintuplet* $[F', \bar{B}, G', \bar{D}; \bar{w}]$ *is a proper backward stochastic realization of* z *with state process* $\bar{x}$ *given by* (2.19) *and state covariance* P^{-1}*, where* $\bar{w}$ *is as in* (2.20) *and* $\bar{B}, \bar{D}$ *are given by* (2.16)-(2.17).

Results closely related to this theorem have been presented by Akaike [1; p.168] and Ruckebusch [33; p.32]. However, the first deals with realizations without noise in the observations, the second does not derive expressions for $\overline{w}$, $\overline{B}$ and $\overline{D}$ such as (2.20), (2.16) and (2.17).

So far we have said nothing about *existence* of proper stochastic realizations. It is well known that a necessary and sufficient condition for a purely nondeterministic wide sense stationary process z to admit finite dimensional stochastic realizations is that its spectral density is rational and that in such a case there exists a unique realization of the type (2.6)-(2.7) corresponding to P_* (cf. [33] for example). The minimum variance realization

$$x_*(t+1) = Fx_*(t) + B_*u_*(t) \tag{2.23}$$

$$z(t) = Hx_*(t) + R(P_*)^{1/2}u_*(t) \tag{2.24}$$

is the *steady-state Kalman filter*, with the *steady-state Kalman gain* B_* given by

$$B_* = (F\Sigma H' + BD')(H\Sigma H' + DD')^{-1/2} = (G - FP_*H')R(P_*)^{-1/2} \tag{2.25}$$

where $[F, B, H, D]$ is any wide sense realization and Σ is the unique nonnegative definite solution to the *Algebraic Riccati Equation*

$$\Sigma = F\Sigma F' - (F\Sigma H' + BD)(H\Sigma H' + DD')^{-1}(H\Sigma F' + DB') + BB' \tag{2.26}$$

The noise u_* is called the *innovation process* and is characterized by the fact that $H_t^-(z) = H_t^-(u_*)$ for all $t \in Z$. Finally, if x is the state process of any proper realization (2.6)-(2.7), we have

$$(2.27) \qquad x_*(t) = \hat{E}\{x(t) \mid H^-_{t-1}(z)\}$$

By duality there exists a proper backward stochastic realization corresponding to $\bar{P}_*$, namely the *backward steady-state Kalman filter*

$$(2.28) \qquad \bar{x}_*(t-1) = F'\bar{x}_*(t) + \bar{B}_*\bar{u}_*(t)$$

$$(2.29) \qquad z(t) = G'\bar{x}_*(t) + \bar{R}(\bar{P}_*)^{1/2}\bar{u}_*(t)$$

Here the *backward steady-state Kalman gain* $\bar{B}_*$ is given by

$$(2.30) \qquad \bar{B}_* = (F'\bar{\Sigma}G + \bar{B}\bar{D}')(G'\bar{\Sigma}G + \bar{D}\bar{D}')^{-1/2} = (H' - F'\bar{P}_*G)\bar{R}(\bar{P}_*)^{-1/2}$$

where $[F', \bar{B}, G', \bar{D}]$ is any backward wide sense realization and $\bar{\Sigma}$ is the unique nonnegative definite solution to the *Dual Algebraic Riccati Equation*

$$(2.31) \qquad \bar{\Sigma} = F'\bar{\Sigma}F - (F'\bar{\Sigma}G + \bar{B}\bar{D}')(G'\Sigma G + \bar{D}\bar{D}')^{-1}(G'\bar{\Sigma}F + \bar{D}\bar{B}') + \bar{B}\bar{B}'$$

The equality $H^+_t(z) = H^+_t(\bar{u}_*)$ for all $t \in Z$ characterizes the *backward innovation process* $\bar{u}_*$. The backward filter satisfies

$$(2.32) \qquad \bar{x}_*(t) = \hat{E}\{\bar{x}(t) \mid H^+_{t+1}(z)\}$$

where $\bar{x}$ is the state of any proper backward realization (2.14)-(2.15). By Theorem 2.5 there exists a proper stochastic realization corresponding to (2.28)-(2.29) (which, as it will be apparent in the next section, is unique)

$$(2.33) \qquad x^*(t+1) = Fx^*(t) + B^*u^*(t)$$

$$(2.34) \qquad z(t) = Hx^*(t) + R(P^*)^{1/2}u^*(t)$$

with state covariance P^*. Then, if x is the state process of any realization,

$$(2.35) \qquad \bar{x}_*(t) = (P^*)^{-1}x^*(t+1) = P^{-1}\hat{E}\{x(t+1)|H^+_{t+1}(z)\}$$

and

$$(2.36) \qquad \hat{E}\{x(t)|H^+_t(z)\} = P(P^*)^{-1}x^*(t)$$

This justifies our choice of working with $P^{-1}x$ rather than x in the backward setting. In fact (2.36) is *not* invariant over $\mathcal{P}$.

Definition 2.6 ([19, 33]). A proper stochastic realization of z with state process x is said to be *internal* if $H(x) \subseteq H(z)$, *external* otherwise.

Internal realizations are of particular interest since they are the only one we can construct from the process z. For example, the minimum and maximum variance realizations introduced in this section are internal. It should be noted that the existence of external realizations depends on H. If $H = H(z)$, for instance, all realizations would be internal.

2.4 Characterization of internal realizations.

Let us consider the spectral representation of z (see e.g. [31]) given by

$$z(t) = \int_{-\pi}^{\pi} e^{i\omega t}\, d\hat{z}(\omega)$$

where $d\hat{z}$ is an orthogonal stochastic measure such that

$$E\{d\hat{z}(\omega)\, d\hat{z}(\omega)^+\} = \Phi(e^{i\omega})\, d\omega$$

(Here $+$ denotes complex conjugation and transposition.) Let $W(z) = H(zI - F)^{-1}B_1 + R(P)^{1/2}$ be a square $(m \times m)$ spectral factor of $\Phi(z)$. Then the process u defined by

$$(2.37) \qquad u(t) = \int_{-\pi}^{\pi} e^{i\omega t}[W(e^{i\omega})]^{-1}\, d\hat{z}(\omega)$$

is a normalized white noise such that $u(t) \in H(z)$ for all t [31; p.41] and consequently $[F, B_1, H, R(P)^{1/2}; u]$ is an internal realization of z. The following result shows that $W(\cdot)$ being a square matrix function is also necessary for a realization to be internal.

Theorem 2.7 ([19, 33]) *A proper stochastic realization is internal if and only if its transfer function is square.*

It follows from this theorem and Remark 2.4 that internal realizations of the form (2.6)-(2.7) are in one to one correspondence with the real symmetric solutions of the matrix equation $\Lambda(P) = 0$. Hence, to characterize further internal realizations, one could derive the discrete time counterpart of the fundamental results of J. C. Willems [40] on the algebraic Riccati equation. However, a result akin to the classification of the solutions fo the algebraic Riccati equation can be obtained directly for the state processes of internal realizations. Notice that once the state $x(t)$ of an internal realization has been determined the input $u(t)$ can be obtained inverting (2.9) as follows

$$u(t) = -R(P)^{-1/2}Hx(t) + R(P)^{-1/2}z(t)$$

(In the case when $R(P)$ is singular we need to perform an appropriate number of differencing operations on the output in various directions (cf. [7] for example) before we can express u in terms of x and z.)

Therefore we turn to the problem of characterizing the state process of internal realizations.

Let us introduce the feedback matrix

$$\Gamma_* = F - B_* R(P_*)^{-1/2} H$$

The matrix Γ_* is asymptotically stable due to the minimality of z [13]. It plays a central role in stochastic realization theory, as it is clear from what follows. In particular we have the following important result, whose continuous time counterpart can be found in [19].

Theorem 2.8 ([33]) *The process* x *is the state of an internal realization if and only if*

$$x(t) = [I - \pi_s]x_*(t) + \pi_s x^*(t) \tag{2.38}$$

where π_s *is the projection onto an invariant subspace* S *of* Γ_* *along* $(P^* - P_*)S^{\perp}$. *The covariance* P *of* x *and* π_s *are related as follows*

$$\pi_s = \pi(P) = (P - P_*)(P^* - P_*)^{-1} \tag{2.39}$$

We shall give a new proof of this theorem, by means of an approach which allows us to characterize also the external realizations in the same framework. Our derivation hinges on the following simple observation. Let $[F, (B_1, B_2), H, R(P)^{1/2}; w]$ be a proper stochastic realization of z with state process x and state covariance P. Then $[\Gamma_*, (B_1 - B_*R(P_*)^{-1/2}R(P)^{1/2}, B_2), R(P_*)^{-1/2}H, R(P_*)^{-1/2}R(P)^{1/2}; w]$ is a proper (nonminimal) stochastic realization of the innovation process u_* with state process $x - x_*$ and state covariance $\tilde{P} = P - P_*$. This can be seen by inverting the filter (2.23)-(2.24) to get

$$(2.40) \qquad x_*(t+1) = \Gamma_* x_*(t) + B_* R(P_*)^{-1/2} z(t)$$

$$(2.41) \qquad u_*(t) = -R(P_*)^{-1/2} H x_*(t) + R(P_*)^{-1/2} z(t)$$

and using (2.6)-(2.7). If we set $\tilde{x}(t) = x(t) - x_*(t)$, we obtain the model

$$(2.42) \qquad \tilde{x}(t+1) = \Gamma_* \tilde{x}(t) + (B_1 - B_* R(P_*)^{-1/2} R(P)^{1/2}) u(t) + B_2 v(t) \qquad w = \begin{pmatrix} u \\ v \end{pmatrix}$$

$$(2.43) \qquad u_*(t) = R(P_*)^{-1/2} H\tilde{x}(t) + R(P_*)^{-1/2} R(P)^{1/2} u(t)$$

which is a forward stochastic realization of u_* since $w(t) \perp H_t^-(\tilde{x})$ for all t. The representation (2.42)-(2.43) is not minimal since u_* is a white noise and its minimal realizations have dimension zero. Conversely consider a forward stochastic realization of u_* of the form

$$(2.44) \qquad \xi(t+1) = \Gamma_* \xi(t) + \tilde{B}_1 u(t) + \tilde{B}_2 v(t) \qquad w = \begin{pmatrix} u \\ v \end{pmatrix}$$

$$(2.45) \qquad u_*(t) = R(P_*)^{-1/2} [H\xi(t) + R(\tilde{P})^{1/2} u(t)]$$

where w is a normalized white noise and $\tilde{B}_1$ is $n \times m$. Observe that $w(t)$ is orthogonal to $H_t^-(x)$, where $x = \xi + x_*$, since $x_*(t) \in H_{t-1}^-(z) = H_{t-1}^-(u_*)$. We conclude from this that $[F, (\tilde{B}_1 + B_* R(\tilde{P})^{1/2}, \tilde{B}_2), H, R(P_*)^{1/2} R(\tilde{P})^{1/2}; w]$ is a minimal stochastic realization of z. We collect these observations in the following

Lemma 2.9 *The map which sends the realization* $[F, (B_1, B_2), H, R(P)^{1/2}; w]$ *to the realization* $[\Gamma_*, (B_1 - B_*R(P_*)^{-1/2}R(P)^{1/2}, B_2), R(P_*)^{-1/2}H, R(P_*)^{-1/2}R(P)^{1/2}; w]$ *is a one to one correspondence between realizations of* z *of the form* (2.6)-(2.7) *and realizations of* u_* *of the form* (2.44)-(2.45).

The map in Lemma 2.9 also induces a correspondence between state covariances which maps $P \in \mathcal{P}$ to $P - P_*$, translating the set $\mathcal{P}$ of the amount $-P_*$. The set $\tilde{\mathcal{P}} = \mathcal{P} - P_*$ has the zero element as its minimum and the positive definite quantity $P^* - P_*$ as its maximum. Notice that the correspondence established in Lemma 2.9 is simply the correspondence between the two input-output relations

$$z(t) = \int_{-\pi}^{\pi} e^{i\omega t} W(e^{i\omega})\, d\hat{w}(\omega)$$

and

$$u_*(t) = \int_{-\pi}^{\pi} e^{i\omega t} W_*^{-1}(e^{i\omega}) W(e^{i\omega})\, d\hat{w}(\omega)$$

where $W(z) = H(zI - F)^{-1}(B_1, B_2) + R(P)^{1/2}$, $d\hat{w}$ is an orthogonal stochastic measure such that $w(t) = \int_{-\pi}^{\pi} e^{i\omega t}\, d\hat{w}(\omega)$ and $W_*(z) = H(zI - F)^{-1}B_* + R(P_*)^{1/2}$.

From (2.23)-(2.24) and (2.40)-(2.41) we know that $H_t^-(z) = H_t^-(u_*)$ for all t and $H(z) = H(u_*)$. Since u_* is a white noise we have the following orthogonal decomposition for the space $H(z)$

$$(2.46) \qquad H(z) = H_{y-1}^-(z) \oplus H_t^+(u_*)$$

Then, if x is the state process of an internal realization, we have

$$x(t) = \hat{E}\{x(t)|H(z)\} = \hat{E}\{x(t)|H_{t-1}^{-}(z)\} + \hat{E}\{x(t)|H_t^{+}(u_*)\}$$

which implies

$$x(t) = x_*(t) + \hat{E}\{x(t) - x_*(t)|H_t^{+}(u_*)\} \tag{2.47}$$

in view of (2.27) and the orthogonality between $x_*(t)$ and $H_t^{+}(u_*)$. To compute $\hat{E}\{x(t) - x_*(t)|H_t^{+}(u_*)\}$ observe first that $\tilde{x}(t) = x(t) - x_*(t)$ is the state process of a realization of u_* of the form (2.41)-(2.42). Secondly, notice that u_* is a stochastic process enjoying all the properties of z. Therefore we simply derive relation (2.36) with $\tilde{x}$ and u_* in place of x and z respectively. This idea of replacing a stochastic process by its innovations is of course very common in filtering theory and it turns out to be helpful also in our context.

We shall now derive the backward counterpart of a realization of the type (2.42)-(2.43) corresponding to an internal realization. We set $B_2 = 0$ in (2.42)-(2.43) and define $\tilde{P} = P - P_*$. An orthogonal decomposition for $\tilde{x}(t)$ as in (2.18) yields the identity

$$\tilde{x}(t) = \tilde{P}\Gamma_*'\tilde{P}^{\#}\tilde{x}(t+1) + [\tilde{x}(t) - \tilde{P}\Gamma_*'\tilde{P}^{\#}\tilde{x}(t+1)] \tag{2.48}$$

Observe that $\tilde{x}(t) - \tilde{P}\Gamma_*'\tilde{P}^{\#}\tilde{x}(t+1)$ is orthogonal to $H_{t-1}^{-}(z)$. Also, using (2.42)-(2.43), we see that $\hat{E}\{\tilde{x}(t) - \tilde{P}\Gamma_*'\tilde{P}^{\#}\tilde{x}(t+1)|H_{t+1}^{+}(u_*)\} = 0$. Hence, using (2.46),

$$\begin{aligned}\tilde{x}(t) - \tilde{P}\Gamma_*'\tilde{P}^{\#}\tilde{x}(t+1) &= \hat{E}\{\tilde{x}(t) - \tilde{P}\Gamma_*'\tilde{P}^{\#}\tilde{x}(t+1)|u_*(t)\} \\ &= \hat{E}\{\tilde{x}(t)|u_*(t)\} = \tilde{P}H'R(P_*)^{-1/2}u_*(t)\end{aligned}$$

and (2.48) becomes

(2.49) $\tilde{x}(t) = \tilde{P}\Gamma_*'\tilde{P}^{\#}\tilde{x}(t+1) + \tilde{P}H'R(P_*)^{-1/2}u_*(t)$

or

(2.50) $\tilde{P}^{\#}\tilde{x}(t) = \tilde{P}^{\#}\tilde{P}\Gamma_*'\tilde{P}^{\#}\tilde{x}(t+1) + \tilde{P}^{\#}\tilde{P}H'R(P_*)^{-1/2}u_*(t)$

The output simply reads

(2.51) $u_*(t) = O\tilde{P}^{\#}\tilde{x}(t+1) + u_*(t)$

where O is the $m \times n$ zero matrix. The model (2.50)-(2.51) is the backward counterpart of (2.42)-(2.43). We stress the fact that all backward realizations of the innovations which we obtain in this fashion from realizations (2.42)-(2.43) with $B_2 = 0$ have the same input noise u_*. For $\tilde{x} = x^* - x_*$ we obtain the backward filter

(2.52)
$$\begin{aligned}(P^* - P_*)^{-1}(x^*(t) - x_*(t)) &= \\ &= \Gamma_*'(P^* - P_*)^{-1}(x^*(t+1) - x_*(t+1)) \\ &\quad + H'R(P_*)^{-1/2}u_*(t)\end{aligned}$$

Using alternatively (2.42) and (2.49) to compute $\hat{E}\{\tilde{x}(t+1)\tilde{x}(t)'\}$ we establish the identity $\Gamma_*\tilde{P} = \tilde{P}\tilde{P}^{\#}\Gamma_*\tilde{P}$ which gives

(2.53) $\tilde{P}\Gamma_*' = \tilde{P}\Gamma_*'\tilde{P}^{\#}\tilde{P}$

Then, using (2.49) and (2.53) we obtain

$$\tilde{x}(t) = \tilde{P}\sum_{i=0}^{\infty}(\Gamma_*')^{i}H'R(P_*)^{-1/2}u_*(t+i)$$

which, together with (2.52), yields the desired expression

(2.54) $x(t) = x_*(t) + (P - P_*)(P^* - P_*)^{-1}(x^*(t) - x_*(t))$

Hence $\tilde{x}(t) \in H(x^*(t) - x_*(t))$ and (2.54) can be written

$$\tilde{x}(t) = \hat{E}\{\tilde{x}(t) | x^*(t) - x_*(t)\} \tag{2.55}$$

$$= (P - P_*)(P^* - P_*)^{-1}(x^*(t) - x_*(t))$$

from which it is seen that $\pi(P) = (P - P_*)(P^* - P_*)^{-1}$ is a projection. Rewriting (2.53) in the form

$$\Gamma_*\pi(P) = \pi(P)(P^* - P_*)\tilde{P}^{\#}\Gamma_*\pi(P)$$

we see that $\pi(P)$ projects onto an invariant subspace of Γ_*. Since $\pi(P)(P^* - P_*) = (P^* - P_*)\pi(P)'$ and $\pi(P)'$ projects along $S^{\perp}$ [15; p.61], we conclude that $\pi(P)$ projects parallel to $(P^* - P_*)S^{\perp}$. Conversely if π projects onto an invariant subspace of Γ_* and $\pi(P^* - P_*) = (P^* - P_*)\pi'$, i.e., π is an *admissable projection* in Ruckebush's language, it is easy to construct first a realization of the innovations and then one (internal) of z along the same lines as in [33]. This completes the proof of Theorem 2.8. //

Remark 2.10. Notice that, given the special form of the realization (2.50)-(2.51), we did not need to invoke any invariance property such as (2.32) of the filter (2.52) to compute $\hat{E}\{\tilde{x}(t) | H_t^+(u_*)\}$. The following interpretation for Theorem 2.8 emerged in the proof. The state process of an internal realization of z is given by the forward filter of z plus a "piece" of the maximum variance error $x^*(t) - x_*(t)$. This piece must be such as to conform with the dynamics of $x^*(t) - x_*(t)$ which is determined by the transition matrix Γ_*, i.e., it must correspond to an invariant subspace of Γ_*.

2.5. External realizations.

It is clear that a necessary condition for the existence of external realizations is the presence in H of elements orthogonal to $H(z)$. For the sake of simplicity we assume that $H = H(z) \oplus H(\zeta)$, where ζ is an n-dimensional normalized white noise orthogonal to $H(z)$. As it will be apparent from what follows, this assumption is the minimum one needed to guarantee the existence of a proper stochastic realization corresponding to each wide sense stochastic realization.

Let x be the state process of a realization (2.6)-(2.7) and P its covariance. Then the counterpart of (2.47) is

$$\tilde{x}(t) = x_*(t) + \hat{E}\{\tilde{x}(t)|H_t^+(u_*)\} + \hat{E}\{\tilde{x}(t)|H(\zeta)\} \tag{2.56}$$

and (2.48) corresponds to

$$\begin{aligned}\tilde{x}(t) = \tilde{P}\Gamma_*'\tilde{P}^{\#}\tilde{x}(t+1) &+ \tilde{P}H'R(P_*)^{-1/2}u_*(t) \\ &+ E\{\tilde{x}(t) - \tilde{P}\Gamma_*'\tilde{P}^{\#}\tilde{x}(t+1)|H(\zeta)\}\end{aligned} \tag{2.57}$$

Now let us assume that ζ is chosen in such a way that the condition $H_{t-1}^-(\zeta) \perp H_t^+(\tilde{x})$ holds and ζ and $\tilde{x}$ are stationarily correlated for every realization (2.42)-(2.43). This assumption is introduced to enable us to treat ζ in the same way as the innovations. It will be clear from what follows that indeed this is a natural assumption when trying to model all realizations using a unique exogenous noise. We can now add to (2.42)-(2.43) the output.

$$\zeta(t) = M\tilde{x}(t) + [\zeta(t) - M\tilde{x}(t)]$$

where $M = \hat{E}\{\zeta(t)\tilde{x}(t)'\}\tilde{P}^{\#}$ and an argument very similar to that used for the innovations gives $\hat{E}\{\tilde{x}(t) - \tilde{P}\Gamma_*'\tilde{P}^{\#}x(t+1) \mid H(\zeta)\} = \tilde{P}M'\zeta(t)$ so that (2.57) becomes

$$\tilde{x}(t) = \tilde{P}\Gamma_*'\tilde{P}^{\#}\tilde{x}(t+1) + \tilde{P}H'R(P_*)^{-1/2}u_*(t) + \tilde{P}M'\zeta(t) \tag{2.58}$$

Note that M must satisfy

$$\tilde{P} = \tilde{P}\Gamma_*'\tilde{P}^{\#}\Gamma_*\tilde{P} + \tilde{P}H'R(P_*)^{-1}H\tilde{P} + \tilde{P}M'M\tilde{P}$$

and that, as in the internal case, the input noise (u_*, ζ) is the same for all realizations.

Let $\tilde{x}_I(t)$ and $\tilde{x}_E(t)$ denote $\hat{E}\{\tilde{x}(t) \mid H_t^+(u_*)\}$ and $\hat{E}\{\tilde{x}(t) \mid H(\zeta)\}$ respectively. Then it follows from (2.58) that

$$\tilde{x}_I(t) = (P - P_*)(P^* - P_*)^{-1}(x^*(t) - x_*(t)) \tag{2.59}$$

and

$$\tilde{x}_E(t) = \tilde{P}\Gamma_*'\tilde{P}^{\#}\tilde{x}_E(t+1) + \tilde{P}M'\zeta(t) \tag{2.60}$$

Using (2.53), (2.56), (2.59) and (2.60) we conclude that

$$x(t) = x_*(t) + (P - P_*)(P^* - P_*)^{-1}(x^*(t) - x_*(t)) + \sum_{i=0}^{\infty} (\Gamma_*')^i M'\zeta(t+i) \tag{2.61}$$

Conversely, given any matrix M such that $M'M \in C_n$, let $\tilde{P}$ solve

$$\tilde{P}^{-1} = \Gamma_*'\tilde{P}^{-1}\Gamma_* + H'R(P_*)^{-1}H + M'M$$

Then, using (2.61), we construct the state of a stochastic realization of z. All the realizations with singular $\tilde{P}$ can be obtained through limiting procedures, using realizations corresponding to unbounded sequences of $M'M$ in the cone C_n.

The derivation of the classification of external realizations presented above is quite similar to the one given in [33; p.65], but we feel it will give some further insight into the concepts described there. Moreover it provides a clear stochastic meaning for the parametric representation of the set P derived by Faurre [12; p.52] in continuous time and by Germain [13; p.61] in discrete time. Finally the input processes of external realizations can be characterized along the same lines as in [19].

Part 3

DISCRETE TIME STOCHASTIC REALIZATION: THE SINGULAR CASE

3.1 Invariant, predictable and smoothable subspaces.

Problems I and II are called *singular* when $\Phi(\infty)$ is singular. It follows from Theorems 1.6 and 1.9 that in the singular case there exist nontrivial invariant directions for the Riccati equation (1.5) associated to every solution to Problem I. Abusing language we shall say that a vector a is invariant (predictable) for $[A, B, C, D]$ if it is invariant (predictable) for the corresponding equation (1.5).

Proposition 3.1 *The space I of invariant directions is invariant over all wide sense realizations of* z.

Proof. Immediate from Theorems 1.6 and 1.9. //

The following result describes the singular case in a number of different ways.

Theorem 3.2 *The following statements are equivalent:*

(i) $\Phi(\infty)$ *is singular.*

(ii) Γ_* *is singular.*

(iii) $R(P^*)$ *is singular.*

(iv) $R(P_*)^{1/2} - B_*'(F')^{-1}H'$ *is singular.*

Proof. Let $\gamma \in R^m$ be in the null space of $\Phi(\infty)$. Then, recalling that $\Phi(\infty) = DD' - DB'(F')^{-1}H'$ where $[F, B, H, D]$ is any wide sense realization, we obtain from (2.25) $B_*'(F')^{-1}H'\gamma = (H\Sigma H' + DD')^{1/2}\gamma = R(P_*)^{1/2}\gamma$. Hence $\gamma \in N(R(P_*)^{1/2} - B_*'(F')^{-1}H')$ and $(F')^{-1}H'\gamma \in N(\Gamma_*')$. Conversely, if (ii) holds, use the fact that the eigenvalues of Γ_* are equal to the zeroes of the determinant of W_* to get (iv) from which (i) follows trivially. The equivalence between (ii) and (iii) has been proven by Ruckebusch [33; p.70]. //

Corollary 3.3 *The set* $\mathcal{Q}$ *is nonempty if and only if* $\Phi(\infty)$ *is singular.*

Proof. For any $P \in \mathcal{Q}$ we have $R(P^*) \le R(P)$. //

This says that the singular case occurs precisely when some of the wide sense realizations have $R(P)$ singular, in particular when $R(P^*)$ is singular. This constrast with the continuous time situation where, when the innovation process is full rank, all the input noises have nonsingular intensity.

Let $T_H(i)$ $i = 0, 1, \ldots$ be as in Theorem 1.9 so that $\Phi(z) = \sum_{i=0}^{\infty} T_H(i)z^{-i}$ for $|z|$ large enough and $\tilde{T}_*$ be the weighting pattern (1.17) corresponding to the minimum variance realization.

Theorem 3.4 *The following statements are equivalent:*

(i) a *is an s-invariant direction of the wide sense realization* $[F, B, H, D]$.

(ii) $a = \sum_{i=1}^{s} (F')^{-i}H'\lambda_i$ *with* $\sum_{i=0}^{s-j} T_H(i)\lambda_{j+i} = 0 \quad j = 1, \ldots, s.$

(iii) $a = \sum_{i=1}^{s} (F')^{-i} H'\lambda_i$ *with* $\sum_{i=0}^{s-j} \tilde{T}_*(i)\lambda_{j+1} = 0 \quad j = 1, \ldots, s.$

(iv) $a = \sum_{i=1}^{s} (F')^{-i} H'\lambda_i$ *with* $a'x_*(t) = \sum_{i=1}^{s} \lambda_i' z(t-i)$ *for all* t.

(v) a *is a generalized eigenvector of rank* s *(an eigenvector if* $s = 1$*) of* Γ_*' *corresponding to the eigenvalue zero.*

Proof. The eigenvalence of (ii) and (iv) is immediate. The rest follows at once from Theorem 1.6, in view of Proposition 3.1 and the fact that the deterministic and stochastic elements in the minimum variance realization can be obtained as limits of the corresponding quantities in a transient Kalman filter of the form (1.3). //

Corollary 3.5 *All the invariant directions of* $[F, B_*, H, R(P_*)^{1/2}]$ *are predictable.*

Proof. It follows directly from Theorem 1.11 and condition (iii) of Theorem 3.4. //

Note that in Theorem 3.4 the space I appears as the invariant subspace of Γ_*' related to the zero eigenvalue. We now introduce the backward counterpart of the concept of invariant direction. A vector $\bar{a}$ is said to be a *dually* s-invariant direction of the dual transient Riccati equation

$$(3.1)\quad \begin{cases} \bar{\Sigma}(t-1) = F'\bar{\Sigma}(t)F - (F'\bar{\Sigma}(t)G + \bar{B}\bar{D}')(G'\bar{\Sigma}(t)G + \bar{D}\bar{D}')^{-1}(G'\bar{\Sigma}(t)F + \bar{D}\bar{B}') \\ \qquad\qquad + \bar{B}\bar{B}' \\ \bar{\Sigma}(0) = \bar{P} \end{cases}$$

if $\bar{a}'\bar{\Sigma}(-t; \bar{p}) = \bar{a}'\bar{\Sigma}(-s; 0)$ for all $t \ge s$ and all $\bar{P} \in C_n$. Also let

$\tilde{\tilde{T}}$ be given by (1.17) with $[F', \overline{B}, G', \overline{R}(\overline{P})^{1/2}]$ in place of $[A, B, C, D]$. Duality now gives the following result.

Corollary 3.6 *The following statements are equivalent:*

(i) $\overline{a}$ *is a dually* s-*invariant direction of the backward wide sense realization* $[F', \overline{B}, G', \overline{D}]$.

(ii) $\overline{a} = \sum_{i=1}^{s} F^{-i}G\mu_i$ *with* $\sum_{i=0}^{s-j} T_H(i)'\mu_{j+i} = 0 \quad j = 1, \ldots, s.$

(iii) $\overline{a} = \sum_{i=1}^{s} F^{-i}G\mu_i$ *with* $\sum_{i=0}^{s-j} \overline{T}_*(i)\mu_{j+i} = 0 \quad j = 1, \ldots, s.$

(iv) $\overline{a} = \sum_{i=1}^{s} F^{-i}G\mu_i$ *with* $\overline{a}'\overline{x}_*(t) = \sum_{i=1}^{s} \mu_i'z(t+i)$ *for all* t.

(v) $\overline{a}$ *is a generalized eigenvector of rank* s *(an eigenvector if* $s = 1$*) of* $\overline{\Gamma}_*' = F' - \overline{B}_*\overline{R}(\overline{P}_*)^{-1/2}G'$ *corresponding to the eigenvalue zero.*

Next we define the dual counterpart of predictability.

Definition 3.7. The n-dimensional vector $\overline{a}$ is called an s-*smoothable direction* of equation (3.1) if

$$\overline{a}'\overline{\Sigma}(-t;\ \overline{P}) = \overline{a}'\overline{\Sigma}(-s;\ \overline{P}) = 0 \quad \text{for all} \quad t \geq s \tag{3.2}$$

The terminology is motivated by the fact that if $\overline{a}$ satisfies (3.2) then, by property (iv) in Corollary 3.6, we can smooth the state of any proper stochastic realization corresponding to $[F', \overline{B}, G', \overline{D}]$ exactly in direction $P^{-1}\overline{a}$. Clearly all the dually invariant directions of $[F', \overline{B}_*, G', \overline{R}(\overline{P}_*)^{1/2}]$ are smoothable. Let $\overline{I}$ indicate the space of the invariant directions of (3.1) which, by Proposition 3.1 and duality, is invariant over all backward wide sense realization. Ruckebusch proved that $\bar{\Gamma}_* = (P^*)^{-1}(P^* - P_*)\Gamma_*'(P^* - P_*)^{-1}P^*$ [33; p.53]. Therefore it follows from Corollary 3.6 that $(P^* - P_*)(P^*)^{-1}\bar{I}$ is the invariant subspace of Γ_* corresponding to the zero eigenvalue. Moreover the dimensions of I and $\bar{I}$ are equal. The following theorem characterizes the predictable subspace of an internal realization and the smoothable subspace of the corresponding backward realization. It also shows that the sum of the dimensions of these two subspaces is constant and equal to $\dim I$.

Theorem 3.8 *Let* x *be the state process of the internal realization* $[F, B_1, H, R(P)^{1/2}; u]$ *and* S *the invariant subspace of* Γ_* *associated with* x *in Theorem* 2.8, *so that* $x(t) = x_*(t) + \pi_s(x^*(t) - x_*(t))$ *with* π_s *given by* (2.39). *Then, if* $a = \sum_{i=1}^{n} (F')^{-i}H'\lambda_i$ *belongs to* $S^\perp \cap I$ *and* $\bar{a} = \sum_{i=1}^{n} F^{-i}G\mu_i$ *belongs to* $P^*(P^* - P_*)^{-1}S \cap \bar{I}$ *we have*

$$a'x(t) = \sum_{i=1}^{n} \lambda_i' z(t - i) \tag{3.3}$$

and

$$\bar{a}'(P^*)^{-1}x(t) = \sum_{i=1}^{n} \mu_i' z(t + i - 1) \tag{3.4}$$

Moreover $\dim(S^\perp \cap I) + \dim(P^*(P^* - P_*)^{-1}S \cap \bar{I}) = \dim I$.

Proof. Since $(P^* - P_*)^{-1}\pi_s(P^* - P_*) = \pi_s'$ and π_s projects parallel to $(P^* - P_*)S^\perp$, we have $a'\pi_s = 0$ and $\bar{a}'(P^*)^{-1}\pi_s = \bar{a}'(P^*)^{-1}$. Properties (iv) of Theorem 3.4 and Corollary 3.6 now yield (3.3) and (3.4) respectively. Let k be the smallest positive integer such that $I = N((\Gamma_*')^k)$; Theorem 3.4(v) insures the existence of such a k. Then we have the direct decomposition $R^n = I \oplus R((\Gamma_*')^k)$, where $R((\Gamma_*')^k)$ is the range space of $(\Gamma_*')^k$, cf. [15; p.166] for example. Consider also the usual orthogonal decomposition $R^n = N((\Gamma_*)^k) \oplus R((\Gamma_*')^k)$, where $N((\Gamma_*)^k) = P^* - P_*)(P^*)^{-1}\bar{I}$. It follows that $\dim(S \cap I) = \dim(S \cap (P^* - P_*)(P^*)^{-1}\bar{I})$. To complete the proof, observe that $I = (I \cap S) \oplus (I \cap S^{\perp})$ and that $\dim(S \cap (P^* - P_*)(P^*)^{-1}\bar{I}) = \dim(P^*(P^* - P_*)^{-1}S \cap \bar{I})$. //

It is worthwhile mentioning that $\bar{a}'(P^*)^{-1}$ in (3.4) has actually the form $\sum_{i=1}^{n} \mu_{n-i}' HF^{i-1}$ with $\sum_{k=0}^{n-j} T(k)'\mu_{n-k} = 0$ for $j = 1, \ldots, n$, as one can readily verify using (2.4)-(2.5) and (2.16)-(2.17) to establish the correspondence between $T(\cdot)'$ and $\tilde{\bar{T}}(\cdot)$. Conversely such a vector leads to a smoothable direction in the backward setting. Hence a predictable-smoothable direction in the forward setting (i.e., a direction in which the state can be computed from a finite number of observations z) has the form $\sum_{i=-n}^{n-1} (F')^i H'\gamma_i$ with $\gamma \in N(\mathsf{T})$, where $\gamma' = (\gamma_{n-1}', \gamma_{n-2}', \ldots, \gamma_o', \gamma_{-n}', \ldots, \gamma_{-1}')$ and T is a block diagonal matrix, the two diagonal blocks being block triangular Toeplitz matrices. The upper one has i^{th} row $[T(i-1)', T(i-2)', \ldots, T(0)', 0, \ldots, 0]$ and the lower one has i^{th} row $[\tilde{T}(i-1), \tilde{T}(i-2), \ldots, \tilde{T}(0), 0, \ldots, 0]$, where $i = 1, \ldots, n$.

The linear hull of the components of $x_*(t)$ and $x^*(t)$ is called the *frame space* [18] and denoted by $H_t^{\square}(z)$. In view of Theorem 2.8, we know that the components of the state at time t of an internal realization belong to $H_t^{\square}(z)$. Let us introduce the subspace $H_{t^+}(z)$ of $H_t^{\square}(z)$, given by the linear hull of elements of the form $a'x_*(t)$ and $\bar{a}'(P^*)^{-1}x^*(t)$, where a varies over I and $\bar{a}$ over $\bar{I}$. By analogy to the continuous time case [10], we shall call $H_{t^+}(z)$ the *germ space*, since it contains linear combinations of differences of the type $\Delta_r z(s) = z(s) - z(s - r)$ and of certain other values of the process z that indicate precisely the degree of "smoothness" of the covariance of z in different directions. Then Theorem 3.8 shows that $\dim(X(t) \cap H_{t^+}(z)) = \dim \mathcal{I}$, where $X(t)$ is the space spanned by the components of the state $x(t)$ of an internal realization. Note that in contrast to the continuous time situation [18], the inclusion $H_{t^+}(z) \subset X(t)$ does not hold. From now on let $\dim \underline{I} = \nu$.

Theorem 3.9 *Let* $[F, B_1, H, R(P)^{1/2}; u]$ *be an internal realization. Then this realization can be embedded in a chain of internal realizations* $[F, B_1(i), H, R(P_i)^{1/2}; u_i]$ *with state spaces* $X_i(t)$, $i = 0, \ldots, \nu$, *such that* $P_o \leq P_1 \leq \cdots \leq P_\nu$, $(X_o(t) \cap H_{t^+}(z)) \subset H^-_{t-1}(z)$ *and* $(X_\nu(t) \cap H_{t^+}(z)) \subset H^+_t(z)$.

Proof. Let S be as in Theorem 3.8 and $a_1, \ldots, a_r$ be a basis for $S^\perp \cap \mathcal{I}$. Then we can generate a family S_i of invariant subspaces of Γ_*, $i = 0, \ldots, \nu$, with $\dim(S_i^\perp \cap \mathcal{I}) = \nu - i$, simply eliminating from $S^\perp$, one at a time, the a_i or adding to $S^\perp$ new linearly independent elements of $\mathcal{I}$, both operations being performed taking due care of the rank of the generalized eigenvectors which are dropped or added, so that the resulting subspace is indeed invariant for Γ_*'. This can be done since $\mathcal{I}$ can be decomposed into cyclic subspaces. Clearly this procedure yields a family of internal realizations which differ only on the germ space and such that $S = S_{\nu-r}$. The state covariances are totally ordered since, if $i < j$ and $x_i(t)$, $x_j(t)$ are the corresponding state processes, $x_i(t)$ is equal to $x_*(t)$ in any direction in which it differs from $x_j(t)$. Finally, by construction, $[F, B_1(0), H, R(P_o)^{1/2}; u_o]$ has a full size predictable subspace and the backward realization corresponding to $[F, B_1(\nu), H, R(P_\nu)^{1/2}; u_\nu]$ has a full size smoothable subspace. Thus, the last assertion of the theorem follows. //

Notice that the chain of realizations in Theorem 3.9 is by no means unique. However the minimum and the maximum realizations are uniquely determined. In the case when Γ_* is cyclic, the number of internal realizations is finite and $\leq 2^n$ [40; Remark 18]. Our work has shown that $2^{n-\nu}(\nu + 1)$ is actually an upper bound in the cyclic case. In fact internal realizations are in one-to-one correspondence with the invariant subspaces of Γ_* and, when Γ_* is cyclic, $\mathcal{I}$ is cyclic and the chain of invariant subspaces constructed in Theorem 3.9 is unique, so that the number of different invariant subspaces of Γ_* is less than or equal to $2^{n-\nu}(\nu + 1)$.

Let us consider a proper external realization of the form (2.6)-(2.7) and an invariant direction $a = \sum_{i=1}^{n} (F')^i H'\lambda_i$ for it which is not predictable. Then two cases can occur. Either $\sum_{i=1}^{n-j} (F')^{-i} H'\lambda_{i+j}$ belongs to $N(B_2)$ for $j = 0, \ldots, n-1$ or it does not. It can be seen that in the first case we are in a situation akin to the one for internal realizations and we can associate to the vector a a smoothable direction in the backward setting. In the second case, which always occurs if $B_2B_2' > 0$, a is invariant but the state cannot be determined exactly from a finite string of observations and we would need to have available the process ζ orthogonal to $H(z)$ and to model external realizations as done in Section 2.5 to be able to calculate the state in ν linearly independent directions. For the sake of brevity, we have avoided here going into details about external realizations. However, it should be clear from our discussion that the sum of the dimensions of the predictable and smoothable subspaces associated with an external realization is less than or equal to ν. This fact has the intuitive meaning of indeterminacy introduced by the presence of the orthogonal component ζ.

The presence of nontrivial invariant directions allows, as it should be expected, for a reduction in the dimension of the filtering algorithms available in the literature. For instance, it is a simple exercise to verify that Faurre's algorithms to compute P_* and P^* [12; p.56] reduce to solving $(n - \varkappa) \times (n - \nu)$ matrix equations, the values of P_* and $(P^*)^{-1}$ on the subspaces I and $\bar{I}$ respectively being known a priori in terms of H, F and G. A similar reduction can be obtained for the fast algorithms which compute the gain (1.4) directly (cf. [17] for example), since it is clear that in an invariant direction the value of the gain can be computed directly in terms of the system matrices.

3.2 Noise free stochastic realization and the singular case.

Akaike, in his important paper [1], deals with Markovian representations of the process z without noise in the output and only in his concluding remarks discusses representations with additive noise terms. Indeed, his work was based on some results of Faurre [11] which, starting from a certain factorization of the covariance matrices, were phrased in terms of noise-free realizations. In subsequent work [12] Faurre turned to a different factorization of the covariance matrices which led naturally to realizations with noise in the output. The same choice has, since then, been made by a number of authors [13, 22, 23, 33], but, up to our knowledge, it has never been explained whether the two approaches are equivalent and, if not, what are the shortcomings of either one. We shall now show that, precisely in the singular case, the first approach presents a considerable disadvantage, in that many minimal Markovian realizations are lost. Let us start considering a minimal factorization (Ξ, Θ, Ψ) (i.e., completely controllable and observable) like the one in [11], namely

$$\Delta_j = E\{z(t+j)z(t)'\} = \Psi\Xi^j\Theta \qquad j = 0, 1, 2, \ldots \tag{3.5}$$

and let $\dim \Xi = r$. On the other hand, since Φ is the double side z-transform of Δ, we have

$$\Delta_j = \begin{cases} HF^{j-1}G & j = 1, 2, 3, \ldots \\ G'(F')^{-1}H' + \Phi(\infty) & j = 0 \end{cases} \tag{3.6}$$

Theorem 3.10 *Let* k *be the dimension of* $N(\Phi(\infty))$ *and assume, without loss of generaliy, that* $\Phi(\infty) = [R'0]$ *where* R *is* $(m-k) \times m$. *Then* (Ξ, Θ, Ψ) *is given, up to a change of basis, by*

$$(\Xi, \Theta, \Psi) = \left(\tilde{F}, \begin{bmatrix} F^{-1}G \\ R \end{bmatrix}, \begin{bmatrix} H & \begin{pmatrix} I \\ 0 \end{pmatrix} \end{bmatrix}\right) \tag{3.7}$$

where

$$\tilde{F} = \begin{bmatrix} F & 0 \\ 0 & 0 \end{bmatrix}$$

the identity matrix is $m - k$ *dimensional and* $r = n + m - k$.

Proof. It is easy to check that the triplet in (3.7) satisfies (3.5).

Also (Ξ, Θ) is controllable and (Ξ, Ψ) is observable. In fact suppose $\alpha = \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix}$ with $\alpha_1 \in R^n$ and $\alpha_2 \in R^{m-k}$ is such that

(3.8) $$(\alpha_1', \alpha_2') \begin{bmatrix} F^{-1}G & G & FG & \cdots & F^{n+m-k-2}G \\ R & 0 & 0 & \cdots & 0 \end{bmatrix} = 0$$

Then we see that α_1 must be zero, which forces $\alpha_2' R = 0$ and finally $\alpha_2 = 0$. We conclude that the controllability matrix in (3.8) is full rank. Similarly the observability matrix is seen to have rank $n + m - k$. The conclusion now follows from the uniqueness, up to an equivalence as in (2.1), of the triplet (Ξ, Θ, Ψ). //

Let us assume for the moment that $\Phi(\infty)$ is nonsingular and consider a proper stochastic realization of z $[F, B, H, D; w]$. Then we can associate to it the noise free model

(3.9) $$\xi(t + 1) = \tilde{F}\xi(t) + \begin{bmatrix} F^{-1}B \\ D - HF^{-1}B \end{bmatrix} \eta(t)$$

(3.10) $$z(t) = [H \; I]\xi(t)$$

where

$$\xi(t) = \begin{pmatrix} F^{-1}x(t + 1) \\ (D - HF^{-1}B)w(t) \end{pmatrix}$$

and $\eta(t) = w(t + 1)$. This induces a one-to-one correspondence between wide sense realizations of the form $[F, B, H, D]$ and noise free wide sense realizations of the form $[\tilde{F}, \chi, (H \; I)]$ which are minimal too in view of Theorem 3.10. If we agree to call realizations $[\tilde{F}, \chi, (H \; I) ; \eta]$

with $\chi(n + m) \times m$ internal, then the above correspondence is one-to-one between internal realizations. In particular it maps $[F, B_*, H, R(P_*)^{1/2}; u_*]$ to a realization related to the *steady state pure filter*, i.e., the second innovation representation IR_2 in Gevers terminology [14].

Suppose now that $\Phi(\infty)$ is as in Theorem 3.10 with $k > 0$. Then it is possible to set up a correspondence similar to the one in the nonsingular case only for a rather small subclass of wide sense realizations. More explicitly, let $[F, B, H, D; w]$ be a realization such that $\tilde{T}(0) = D' - B'(F')^{-1}H'$ has rank $m - k$ and V an orthogonal matrix such that $[D - HF^{-1}B]V = \begin{bmatrix} S \\ 0 \end{bmatrix}$ where S is $(m - k) \times p$, p being the number of columns of B. Then we have the $n + m - k$ dimensional noise free model

$$(3.11) \qquad \xi(t + 1) = \tilde{F}\xi(t) + \begin{bmatrix} F^{-1}B \\ S \end{bmatrix}\eta(t)$$

$$(3.12) \qquad z(t) = \begin{bmatrix} H & \begin{pmatrix} I \\ 0 \end{pmatrix} \end{bmatrix}\xi(t)$$

where $\xi(t) = \begin{pmatrix} F^{-1}x(t + 1) \\ SV'w(t) \end{pmatrix}$ and $\eta(t) = V'w(t + 1)$. The wide sense realization given by (3.11)-(3.12) is minimal. This establishes a one-to-one correspondence between minimal wide sense realizations of z such that $\tilde{T}(0)$ has rank $m - k$ and minimal wide sense realizations of the form $[\tilde{F}, \chi, (H \begin{pmatrix} I \\ 0 \end{pmatrix})]$. It is now apparent that the choice of seeking noise free representation of z can cost us, in the singular case, the loss of a considerable number of realizations. Indeed, it is not hard to see that the subset of $\mathcal{P}$ corresponding to realizations with rank $\tilde{T}(0) = m - k$ lie, as $\mathcal{Q}$, in the relative boundary of $\mathcal{P}$.

This shows that, in discrete time, the factorization (3.6) and the associated choice of $H_{t-1}^{-}(z)$, instead of $H_{t}^{-}(z)$, as past space at time t, is more convenient, even though it implies the unpleasant fact that white noise processes have zero dimensional minimal realizations.

Acknowledgments

The author would like to thank his dissertation advisor Professor A. Lindquist for his constant encouragement and many valuable suggestions and Professor G. Picci with whom he had countless discussions on the stochastic realization problem which led to a considerable improvement of this paper.

REFERENCES

1. H. Akaike, *Markovian representation of stochastic processes by canonical variables*, SIAM J. Control 13(1975), 162-173.

2. ________, *Stochastic theory of minimal realization*, IEEE Trans. Automatic Control AC-19(1974), 667-674.

3. B. D. O. Anderson, *Dual form of a positive real lemma*, Proc. IEEE 55(1967), 1749-1750.

4. ________________, *The inverse problem of stationary covariance generation*, J. Statistical Physics 1(1969), 133-147.

5. B. D. O. Anderson and L. Hitz, *Discrete positive-real functions and their applications to system stability*, Proc. IEE.116(1969), 153-155.

6. R. W. Brockett, *Finite Dimensional Linear Systems*, Wiley, 1970.

7. A. E. Bryson and Y. C. Ho, *Applied Optimal Control*, Blaisdell, 1969.

8. R. S. Bucy, D. Rappaport and L. M. Silverman, *Correlated noise filtering and invariant directions for the Riccati equation*, IEEE Trans. Automatic Control AC-15(1970), 535-540.

9. D. J. Clements and B. D. O. Anderson, *Linear-quadratic discrete-time control and constant directions*, Automatica 13(1977), 255-264.

10. H. Dym and H. P. McKean, *Gaussian Processes, Function Theory, and the Inverse Spectral Problem*, Academic Press, 1976.

11. P. Faurre, *Identification par minimisation d'une représentation Markovienne de processus aléatoire*, Symposium on Optimization, Lecture Notes in Mathematics 132, Springer, Berlin, 1970, 83-107.

12. ________, *Réalisations markoviennes de processus stationnaires*, Research Report no.13, March 1973, IRIA(LABORIA), Le Chesnay, France.

13. F. Germain, *Algoritmes de calcul de réalisations markoviennes. Cas singuliers et stabilité*, Research Report no.66, April 1974, IRIA (LABORIA), Le Chesnay, France.

14. M. R. Gevers, *Structural properties of realizations of discrete-time Markovian processes*, Technical Report no. 7050-19, May 1972, Information Systems Laboratory, Stanford University, Stanford, California.

15. P. R. Halmos, *Finite Dimensional Vector Spaces*, Princeton University Press, 1942.

16. B. L. Ho and R. E. Kalman, *Effective construction of linear state-variable models from input/output functions*, Proc. Third Allerton Conf. (1965), 449-459.

17. A. Lindquist, *A new algorithm for optimal filtering of discrete-time stationary processes*, SIAM J. Control 12(1974), 736-746.

18. A. Lindquist and G. Picci, *On the structure of minimal splitting subspaces in stochastic realization theory*, Proc. 1977 Decision and Control Conference, New Orleans, 42-48.

19. ______________________, *On the stochastic realization problem*, SIAM J. Control and Optimization (forthcoming).

20. ______________________, *A state space theory for stationary stochastic processes*, Proc. 21st. Midwest Symposium on Circuits and Systems, August 1978, Aimes.

21. ______________________, *A Hardy space approach to the stochastic realization problem*, to appear in Proc. 1978 Decision and Control Conf., San Diego.

22. A. Lindquist, G. Picci and G. Ruckebusch, *On minimal splitting subspaces and Markovian representations*, to appear.

23. M. Pavon, *Constant directions and singular stochastic realization: the scalar case*, Proc. 1re Colloque AFCET-SMF, Sept. 1978, Palaiseau, France.

24. M. Pavon and R. J.-B. Wets, *A stochastic variational approach to the duality between estimation and control: discrete time*, Proc. the Analysis and Optimization of Stochastic Systems Conf., Sept. 1978, Oxford, Great Britain.

25. H. J. Payne and L. M. Silverman, *On the discrete time algebraic Riccati equation*, IEEE Trans. Automatic Control AC 18(1973), 226-234.

26. R. Penrose, *A generalized inverse for matrices*, Proc. Cambridge Phil. Soc. 51(1955), 406-413.

27. G. Picci, *Stochastic realization of gaussian processes*, Proc. IEEE 64(1976), 112-122.

28. V. M. Popov, *Hyperstability of Control Systems*, Springer-Verlag, 1973.

29. D. Rappaport, *Constant directions of the Riccati equation*, Automatica 8(1972), 175-186.

30. D. Rappaport and L. M. Silverman, *Structure and stability of discrete-time optimal systems*, IEEE Trans. Automatic Control AC-16(1971), 227-232.

31. Y. A. Rozanov, *Stationary Random Processes*, Holden-Day, 1967.

32. ____________, *On Markov extensions of a random process*, SIAM Theory of Probability 22-1(1977), 190-195.

33. G. Ruckebusch, *Représentations Markoviennes de processus Gaussien Stationnaires*, Thèse 3ème cycle, Univ. of Paris VI, May, 1975.

34. ____________, *Représentations Markoviennes de processus Gaussien stationnaires et applications statistiques*, Grenoble, June 1977, Springer-Verlag Lecture Notes in Mathematics 636.

35. ____________, *A state space approach to the stochastic realization problem*, Proc. 1978 Intern. Symp. Circuits and Systems, New York.

36. ____________, *Factorizations minimales de densités spectrales et représentations Markoviennes*, Proc. 1re Colloque AFCET-SMF, Sept. 1978, Palaiseau, France.

37. G. S. Sidhu and U. B. Desay, *New smoothing algorithms based on reverse-time lumped models*, IEEE Trans. Automatic Control AC-21 (1976), 538-541.

38. L. M. Silverman, *Discrete Riccati equations: alternative algorithms, asymptotic properties and system theory interpretations*, Control and Dynamic Systems 12(1976), 313-386.

39. L. M. Silverman and H. E. Meadows, *Equivalence and synthesis of time variable linear systems*, Proc. Fourth Allerton Conf. (1966), 776-784.

40. J. C. Willems, *Least square stationary optimal control and the algebraic Riccati equation*, IEEE Trans. Automatic Control AC-16 (1971), 621-634.

41. D. C. Youla and P. Tissi, *n-port synthesis via reactance extraction--Part I*, IEEE Intern. Convention Record 14(1966), 183-205.

Unclassified
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

19 REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 18 AFOSR-TR-79-0807	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) 6 STOCHASTIC REALIZATION AND INVARIANT DIRECTIONS OF THE MATRIX RICCATI EQUATION		5. TYPE OF REPORT & PERIOD COVERED 9 Interim rept.
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) 10 Michele/Pavon		8. CONTRACT OR GRANT NUMBER(s) 15 AFOSR-78-3519, NSF-MPS-77-00241
9. PERFORMING ORGANIZATION NAME AND ADDRESS University of Kentucky Department of Mathematics Lexington, KY 40506		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 61102F 16 2304/A1 17
11. CONTROLLING OFFICE NAME AND ADDRESS Air Force Office Of Scientific Research/NM Bolling AFB, Washington, D.C. 20332		12. REPORT DATE 11 1979
		13. NUMBER OF PAGES 60
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office) 12 62p.		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION ST. ENT (of abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY TES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Invariant directions of the Riccati difference equation of Kalman filtering are shown to occur in a large class of prediction problems and to be related to a certain invariant subspace of the transpose of the feedback matrix. The discrete time stochastic realization problem is studied in its deterministic as well as probabilistic aspects. In particular a new derivation of the classification of the minimal markovain representations of the given process z is presented which is based on a certain backward filter of the innovations. For each markovian representation which can be dtermined from z the space of invariant directio-
(continued)

DD FORM 1 JAN 73 1473

UNCLASSIFIED 410 906

20 Abstract (continued)

is decomposed into two subspaces, one on which it is possible to predict the state process without error forward in time and one on which this can be done backward in time.